

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 22:04:37 : Search time 37.7391 Seconds
(without alignments)
27.692 Million cell updates/sec

Title: US-10-634-914-16

Perfect score: 76

Sequence: 1 ETEPLPKGVTPRS 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	259 3	US-08-809-326A-5	Sequence 5, Appli
2	76	100.0	259 4	US-09-689-914A-5	Sequence 5, Appli
3	76	100.0	259 4	US-09-689-913A-5	Sequence 5, Appli
4	76	100.0	259 4	US-09-689-916A-5	Sequence 5, Appli
5	76	100.0	271 3	US-08-809-326A-2	Sequence 2, Appli
6	76	100.0	271 4	US-09-689-914A-2	Sequence 2, Appli
7	76	100.0	271 4	US-09-689-913A-2	Sequence 2, Appli
8	76	100.0	271 4	US-09-689-916A-2	Sequence 2, Appli
9	76	100.0	432 3	US-08-809-326A-16	Sequence 16, Appli
10	76	100.0	432 4	US-09-689-914A-16	Sequence 16, Appli
11	76	100.0	432 4	US-09-689-913A-16	Sequence 16, Appli
12	76	100.0	432 4	US-09-689-916A-16	Sequence 16, Appli
13	76	100.0	488 3	US-08-809-326A-1	Sequence 1, Appli
14	76	100.0	488 4	US-09-689-914A-1	Sequence 1, Appli
15	76	100.0	488 4	US-09-689-913A-1	Sequence 1, Appli
16	76	100.0	488 4	US-09-689-916A-1	Sequence 1, Appli
17	76	100.0	496 4	US-09-198-452A-867	Sequence 867, App
18	76	100.0	496 4	US-09-438-185A-811	Sequence 811, App
19	76	100.0	649 3	US-08-809-326A-15	Sequence 15, Appli
20	76	100.0	649 4	US-09-689-914A-15	Sequence 15, Appli
21	76	100.0	649 4	US-09-689-913A-15	Sequence 15, Appli
22	76	100.0	649 4	US-09-689-916A-15	Sequence 15, Appli
23	46	60.5	243 4	US-09-270-767-32265	Sequence 32265, A
24	46	60.5	243 4	US-09-270-767-47482	Sequence 47482, A
25	44	57.9	175 4	US-09-902-540-10741	Sequence 10741, A
26	44	57.9	671 4	US-09-252-991A-16922	Sequence 16922, A
27	44	57.9	1257 2	US-08-750-152A-2	Sequence 2, Appli

28	43	56.6	236 4	US-09-252-991A-26874	Sequence 26874, A
29	43	56.6	261 4	US-09-252-991A-27499	Sequence 27499, A
30	43	56.6	731 4	US-09-252-991A-17180	Sequence 17180, A
31	42	55.3	159 4	US-09-252-991A-29789	Sequence 29789, A
32	42	55.3	247 2	US-08-951-822-23	Sequence 23, Appli
33	42	55.3	247 3	US-08-705-245-4	Sequence 23, Appli
34	42	55.3	247 3	US-09-368-951-23	Sequence 23, Appli
35	42	55.3	247 4	US-09-390-207-16	Sequence 16, Appli
36	42	55.3	247 4	US-09-229-947-23	Sequence 23, Appli
37	42	55.3	247 4	US-09-490-714-4	Sequence 23, Appli
38	42	55.3	247 4	US-09-949-016-9554	Sequence 9554, Ap
39	42	55.3	252 1	US-08-462-169B-2	Sequence 2, Appli
40	42	55.3	252 1	US-08-462-169B-23	Sequence 23, Appli
41	42	55.3	252 3	US-09-103-079-2	Sequence 23, Appli
42	42	55.3	252 3	US-09-103-079-23	Sequence 23, Appli
43	42	55.3	252 4	US-09-425-021-2	Sequence 2, Appli
44	42	55.3	252 4	US-09-425-021-23	Sequence 23, Appli
45	42	55.3	252 4	US-09-564-829-2	Sequence 2, Appli
46	42	55.3	278 4	US-09-252-991A-28712	Sequence 28712, A
47	42	55.3	503 4	US-09-252-991A-27888	Sequence 27888, A
48	42	55.3	1164 3	US-08-923-992A-2	Sequence 2, Appli
49	42	55.3	1479 3	US-08-840-062-2	Sequence 2, Appli
50	42	55.3	1479 3	US-08-840-062-4	Sequence 4, Appli
51	41.5	54.6	1023 4	US-10-164-595-20	Sequence 20, Appli
52	41.5	54.6	1070 4	US-10-164-595-18	Sequence 18, Appli
53	41.5	54.6	1073 4	US-10-164-595-22	Sequence 22, Appli
54	41	53.9	139 4	US-09-198-452A-1197	Sequence 1197, Ap
55	41	53.9	224 4	US-09-252-991A-24969	Sequence 24969, A
56	41	53.9	239 4	US-09-543-681A-4932	Sequence 4932, Ap
57	41	53.9	292 4	US-09-270-767-44114	Sequence 44114, A
58	41	53.9	307 4	US-09-252-991A-21435	Sequence 21435, A
59	41	53.9	346 4	US-09-252-991A-22386	Sequence 22386, A
60	41	53.9	393 4	US-09-463-158A-2	Sequence 2, Appli
61	41	53.9	447 3	US-09-199-637A-351	Sequence 351, App
62	41	53.9	559 4	US-10-116-370-2	Sequence 2, Appli
63	41	53.9	567 4	US-09-949-016-10952	Sequence 10952, A
64	41	53.9	582 3	US-09-428-711A-2	Sequence 2, Appli
65	41	53.9	593 4	US-09-252-991A-23251	Sequence 23251, A
66	41	53.9	904 4	US-09-976-594-615	Sequence 615, App
67	41	53.9	1045 4	US-09-438-185A-485	Sequence 485, App
68	40.5	53.3	331 2	US-08-997-080-182	Sequence 182, App
69	40.5	53.3	331 2	US-08-997-362-182	Sequence 182, App
70	40.5	53.3	331 3	US-09-095-855-182	Sequence 182, App
71	40.5	53.3	331 3	US-09-324-542-182	Sequence 182, App
72	40.5	53.3	331 4	US-09-205-426-182	Sequence 182, App
73	40	52.6	75 3	US-08-865-297-7	Sequence 7, Appli
74	40	52.6	87 3	US-08-776-971-59	Sequence 59, Appli
75	40	52.6	87 3	US-08-776-971-135	Sequence 135, App
76	40	52.6	87 3	US-08-776-971-138	Sequence 138, App
77	40	52.6	87 4	US-09-576-290-59	Sequence 59, Appli
78	40	52.6	87 4	US-09-576-290-135	Sequence 135, App
79	40	52.6	87 4	US-09-576-290-138	Sequence 138, App
80	40	52.6	184 4	US-09-252-991A-23404	Sequence 23404, A
81	40	52.6	189 3	US-08-865-297-2	Sequence 2, Appli
82	40	52.6	270 4	US-09-543-681A-8304	Sequence 8304, Ap
83	40	52.6	336 4	US-09-270-767-44787	Sequence 44787, A
84	40	52.6	342 4	US-09-270-767-32874	Sequence 32874, A
85	40	52.6	342 4	US-09-270-767-48091	Sequence 48091, A
86	40	52.6	374 3	US-09-091-405-2	Sequence 2, Appli
87	40	52.6	415 2	US-08-833-642A-5	Sequence 5, Appli
88	40	52.6	419 4	US-09-248-796A-21670	Sequence 21670, A
89	40	52.6	520 4	US-09-266-965-108	Sequence 108, App
90	40	52.6	631 3	US-08-906-360-1	Sequence 1, Appli
91	40	52.6	659 4	US-09-147-119-7	Sequence 7, Appli
92	40	52.6	707 4	US-09-949-016-8495	Sequence 8495, Ap
93	40	52.6	728 4	US-09-508-824-10	Sequence 10, Appli
94	40	52.0	69 4	US-09-513-999C-4418	Sequence 4418, Ap
95	39.5	52.0	525 4	US-09-107-532A-5095	Sequence 5095, Ap
96	39.5	51.3	22 2	US-08-769-745-41	Sequence 41, Appli
97	39	51.3	25 2	US-08-769-745-8	Sequence 8, Appli
98	39	51.3	115 4	US-09-252-991A-25641	Sequence 25641, A
99	39	51.3	122 3	US-09-247-352-2	Sequence 2, Appli
100	39	51.3			

ALIGNMENTS

RESULT 1
US-08-809-326A-5
; Sequence 5, Application US/08809326A
; Patent No. 6165478
; GENERAL INFORMATION:
; APPLICANT: Izutsu, Hiroshi
; APPLICANT: Obata, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,326A
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 224711/94
; FILING DATE: 20-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106006/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106008/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106009/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106010/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106011/95
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Miller, Charles E.
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 7426-043-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 259 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-809-326A-5

Query Match 100.0%; Score 76; DB 3; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||
Db 164 ETPELPKPGVTPRS 177

RESULT 2

US-09-689-914A-5
; Sequence 5, Application US/09689914A
; Patent No. 6485914
; GENERAL INFORMATION:
; APPLICANT: Hiroshi Izutsu et al.
; TITLE OF INVENTION: METHOD FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: 1254-0166P
; CURRENT APPLICATION NUMBER: US/09/689,914A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Chlamydoiphila pneumoniae
US-09-689-914A-5

Query Match 100.0%; Score 76; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||
Db 164 ETPELPKPGVTPRS 177

RESULT 3

US-09-689-913A-5
; Sequence 5, Application US/09689913A
; Patent No. 6489122
; GENERAL INFORMATION:
; APPLICANT: Hiroshi Izutsu et al.
; TITLE OF INVENTION: METHOD OF DETECTING ANTI-CHLAMYDIA PNEUMONIAE ANTIBODY USING
; FILE REFERENCE: 1254-0168P
; CURRENT APPLICATION NUMBER: US/09/689,913A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Chlamydoiphila pneumoniae
US-09-689-913A-5

Query Match 100.0%; Score 76; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||
Db 164 ETPELPKPGVTPRS 177

RESULT 4

US-09-689-916A-5
; Sequence 5, Application US/09689916A
; Patent No. 6491924
; GENERAL INFORMATION:
; APPLICANT: Hiroshi Izutsu et al.
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDE
; FILE REFERENCE: 1254-0167P

; CURRENT APPLICATION NUMBER: US/09/689,916A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Chlamydomophila pneumoniae
; US-09-689-916A-5

Query Match 100.0%; Score 76; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPRS 14
Db 164 ETPELPKPGVTPRS 177

RESULT 5

US-08-809-326A-2
; Sequence 2, Application US/08809326A
; Patent No. 6165478

; GENERAL INFORMATION:

; APPLICANT: Izutsu, Hiroshi
; APPLICANT: Obara, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
; TITLE OF INVENTION: ANTIBODY METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,326A
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 224711/94
; FILING DATE: 20-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106006/95
; FILING DATE: 28-APR-1995
; APPLICATION NUMBER: JP 106008/95
; FILING DATE: 28-APR-1995
; APPLICATION NUMBER: JP 106009/95
; FILING DATE: 28-APR-1995
; APPLICATION DATA:
; APPLICATION NUMBER: JP 106010/95
; FILING DATE: 28-APR-1995
; APPLICATION DATA:
; APPLICATION NUMBER: JP 106011/95
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:

; NAME: Miller, Charles E.
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 7426-043-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-809-326A-2

Query Match 100.0%; Score 76; DB 3; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPRS 14
Db 164 ETPELPKPGVTPRS 177

RESULT 6

US-09-689-914A-2
; Sequence 2, Application US/09689914A
; Patent No. 6485914

; GENERAL INFORMATION:

; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: 1254-0166P
; CURRENT APPLICATION NUMBER: US/09/689,914A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 271

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: fusion peptide
; US-09-689-914A-2

Query Match 100.0%; Score 76; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPRS 14
Db 164 ETPELPKPGVTPRS 177

RESULT 7

US-09-689-913A-2
; Sequence 2, Application US/09689913A
; Patent No. 6489122

; GENERAL INFORMATION:

; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD OF DETECTING ANTI-CHLAMYDIA PNEUMONIAE ANTIBODY USING
; FILE REFERENCE: 1254-0168P
; CURRENT APPLICATION NUMBER: US/09/689,913A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 271

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: fusion peptide
; US-09-689-913A-2

Query Match 100.0%; Score 76; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||||
Db 164 ETPELPKPGVTPRS 177

RESULT 8

US-09-689-916A-2
; Sequence 2, Application US/09689916A
; Patent No. 6491924
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDE
; FILE REFERENCE: 1254-0167P
; CURRENT APPLICATION NUMBER: US/09/689,916A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion peptide
US-09-689-916A-2

Query Match 100.0%; Score 76; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||||
Db 164 ETPELPKPGVTPRS 177

RESULT 9

US-08-809-326A-16
; Sequence 16, Application US/08809326A
; Patent No. 6165478

; GENERAL INFORMATION:
; APPLICANT: Izutsu, Hiroshi
; APPLICANT: Obara, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,326A
; FILING DATE: 19-MAR-1997

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 224711/94
; FILING DATE: 20-SEP-1994

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106006/95
; FILING DATE: 28-APR-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106008/95
; FILING DATE: 28-APR-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106009/95
; FILING DATE: 28-APR-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106010/95
; FILING DATE: 28-APR-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106011/95
; FILING DATE: 28-APR-1995

; ATTORNEY/AGENT INFORMATION:
; NAME: Miller, Charles E.
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 7426-043-999

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741

; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 432 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-809-326A-16

Query Match 100.0%; Score 76; DB 3; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||||
Db 325 ETPELPKPGVTPRS 338

RESULT 10

US-09-689-914A-16
; Sequence 16, Application US/09689914A
; Patent No. 6485914

; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: 1254-0166P
; CURRENT APPLICATION NUMBER: US/09/689,914A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion peptide
US-09-689-914A-16

Query Match 100.0%; Score 76; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
|||||
Db 325 ETPELPKPGVTPRS 338


```
RESULT 11
US-09-689-913A-16
; Sequence 16, Application US/09689913A
; Patent No. 6489122
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD OF DETECTING ANTI-CHLAMYDIA PNEUMONIAE ANTIBODY USING
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE-SPECIFIC ANTIGENS
; FILE REFERENCE: 1254-0168P
; CURRENT APPLICATION NUMBER: US/09/689,913A
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion peptide
US-09-689-913A-16

Query Match          100.0%; Score 76; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||
Db      325 ETPELPKPGVTPRS 338

RESULT 12
US-09-689-916A-16
; Sequence 16, Application US/09689916A
; Patent No. 6491924
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDE
; FILE REFERENCE: 1254-0167P
; CURRENT APPLICATION NUMBER: US/09/689,916A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion peptide
US-09-689-916A-16

Query Match          100.0%; Score 76; DB 4; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||
Db      325 ETPELPKPGVTPRS 338

RESULT 13
US-08-809-326A-1
; Sequence 1, Application US/08809326A
; Patent No. 6165478
; GENERAL INFORMATION:
; APPLICANT: Izutsu, Hiroshi
; APPLICANT: Obara, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
```

```
;; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/O
;; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
;; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
;; TITLE OF INVENTION: PNEUMONIAE GENE
;; NUMBER OF SEQUENCES: 31
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Pennie & Edmonds LLP
;; STREET: 1155 Avenue of the Americas
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10036-2711
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/809,326A
;; FILING DATE: 19-MAR-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 224711/94
;; FILING DATE: 20-SEP-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 106006/95
;; FILING DATE: 28-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 106008/95
;; FILING DATE: 28-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 106009/95
;; FILING DATE: 28-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 106010/95
;; FILING DATE: 28-APR-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 106011/95
;; FILING DATE: 28-APR-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Miller, Charles E.
;; REGISTRATION NUMBER: 24,576
;; REFERENCE/DOCKET NUMBER: 7426-043-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212) 790-9090
;; TELEFAX: (212) 869-8864/9741
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 488 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-809-326A-1

Query Match          100.0%; Score 76; DB 3; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||
Db      164 ETPELPKPGVTPRS 177

RESULT 14
US-09-689-914A-1
; Sequence 1, Application US/09689914A
; Patent No. 6485914
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: 1254-0166P
; CURRENT APPLICATION NUMBER: US/09/689,914A
```

```

; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Chlamydophila pneumoniae
US-09-689-914A-1
```

```

Query Match          100.0%; Score 76; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTPRS 14
        |||
Db       164 ETPELPKPGVTPRS 177
```

RESULT 15

```

US-09-689-913A-1
; Sequence 1, Application US/09689913A
; Patent No. 6489122
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD OF DETECTING ANTI-CHLAMYDIA PNEUMONIAE ANTIBODY USING
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE-SPECIFIC ANTIGENS
; FILE REFERENCE: 1254-0168P
; CURRENT APPLICATION NUMBER: US/09/689,913A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Chlamydophila pneumoniae
US-09-689-913A-1
```

```

Query Match          100.0%; Score 76; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTPRS 14
        |||
Db       164 ETPELPKPGVTPRS 177
```

RESULT 16

```

US-09-689-916A-1
; Sequence 1, Application US/09689916A
; Patent No. 6491924
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDE
; FILE REFERENCE: 1254-0167P
; CURRENT APPLICATION NUMBER: US/09/689,916A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Chlamydophila pneumoniae
US-09-689-916A-1
```

```

Query Match          100.0%; Score 76; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTPRS 14
        |||
Db       164 ETPELPKPGVTPRS 177
```

```

RESULT 17
US-09-198-452A-867
; Sequence 867, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 867
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-867
```

```

Query Match          100.0%; Score 76; DB 4; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTPRS 14
        |||
Db       167 ETPELPKPGVTPRS 180
```

RESULT 18

```

US-09-438-185A-811
; Sequence 811, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 811
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0809
US-09-438-185A-811
```

```

Query Match          100.0%; Score 76; DB 4; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTPRS 14
        |||
Db       167 ETPELPKPGVTPRS 180
```

RESULT 19

```

US-08-809-326A-15
; Sequence 15, Application US/08809326A
; Patent No. 6165478
; GENERAL INFORMATION:
; APPLICANT: Izutsu, Hiroshi
; APPLICANT: Obata, Kazuhiko
; APPLICANT: Matsumoto, Akira
```

```
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE GENE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/809,326A
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 224711/94
; FILING DATE: 20-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106006/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106008/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106009/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106010/95
; FILING DATE: 28-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 106011/95
; FILING DATE: 28-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Miller, Charles E.
; REGISTRATION NUMBER: 24,576
; REFERENCE/DOCKET NUMBER: 7426-043-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 649 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-809-326A-15

Query Match 100.0%; Score 76; DB 3; length 649;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 325 ETPELPKPGVTPRS 338
```

```
RESULT 20
US-09-689-914A-15
; Sequence 15, Application US/09689914A
```

```
; Patent No. 6485914
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE
; FILE REFERENCE: 1254-0166P
; CURRENT APPLICATION NUMBER: US/09/689,914A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion peptide
; US-09-689-914A-15

Query Match 100.0%; Score 76; DB 4; length 649;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 325 ETPELPKPGVTPRS 338
```

```
RESULT 21
US-09-689-913A-15
; Sequence 15, Application US/09689913A
; Patent No. 6489122
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: METHOD OF DETECTING ANTI-CHLAMYDIA PNEUMONIAE ANTIBODY USING
; FILE REFERENCE: 1254-0168P
; CURRENT APPLICATION NUMBER: US/09/689,913A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: fusion peptide
; US-09-689-913A-15
```

```
Query Match 100.0%; Score 76; DB 4; length 649;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 325 ETPELPKPGVTPRS 338
```

```
RESULT 22
US-09-689-916A-15
; Sequence 15, Application US/09689916A
; Patent No. 6491924
; GENERAL INFORMATION:
; APPLICANT: Hiroshi IZUTSU et al.
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDE
; FILE REFERENCE: 1254-0167P
; CURRENT APPLICATION NUMBER: US/09/689,916A
; CURRENT FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 649
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

; OTHER INFORMATION: fusion peptide
US-09-689-916A-15

Query Match 100.0%; Score 76; DB 4; Length 649;
Best Local Similarity 100.0%; Pred. No. 0.0052;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETELPKPGVTPRS 14
Db 325 ETELPKPGVTPRS 338

RESULT 23

US-09-270-767-32265
; Sequence 32265, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 32265
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-32265

Query Match 60.5%; Score 46; DB 4; Length 243;
Best Local Similarity 80.0%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVTP 12
Db 153 PELPKPKLTP 162

RESULT 24

US-09-270-767-47482
; Sequence 47482, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47482
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-47482

Query Match 60.5%; Score 46; DB 4; Length 243;
Best Local Similarity 80.0%; Pred. No. 33;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVTP 12
Db 153 PELPKPKLTP 162

RESULT 25

US-09-902-540-10741
; Sequence 10741, Application US/09902540

; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10741
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10741

Query Match 57.9%; Score 44; DB 4; Length 175;
Best Local Similarity 61.5%; Pred. No. 45;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ETELPKPGVTPR 13
Db 58 EEPEVPPPVVPR 70

RESULT 26

US-09-252-991A-16922
; Sequence 16922, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16922
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16922

Query Match 57.9%; Score 44; DB 4; Length 671;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 PELPKPGVTPRS 14
Db 632 PRLPKPDARPRS 643

RESULT 27

US-08-750-152A-2
; Sequence 2, Application US/08750152A
; Patent No. 5977331
; GENERAL INFORMATION:
; APPLICANT: ASAKURA, YOKO
; APPLICANT: KIMURA, EIICHIRO
; APPLICANT: ABE, CHIZU
; APPLICANT: KAWAHARA, YOSHIO
; APPLICANT: NAKAMATSU, TSUYOSHI
; TITLE OF INVENTION: ALPHA-KETOGLUTARATE DEHYDROGENASE GENE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, WAIER & NEUSTADT, P.C.

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22152
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,152A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1257 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-750-152A-2

Query Match 57.9%; Score 44; DB 2; Length 1257;
Best Local Similarity 58.3%; Pred. No. 3.5e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETELPKPGVTP 12
:|:|:|:|:|
Db 131 QQPKLPEPGTP 142

RESULT 28
US-09-252-991A-26874
; Sequence 26874, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26874
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26874

Query Match 56.6%; Score 43; DB 4; Length 236;
Best Local Similarity 66.7%; Pred. No. 85;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 PELPKPGVTPRS 14
|||:|:|:|:|
Db 199 PELPFGVTGKT 210

RESULT 29
US-09-252-991A-27499

; Sequence 27499, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27499
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27499

Query Match 56.6%; Score 43; DB 4; Length 261;
Best Local Similarity 58.3%; Pred. No. 94;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 PELPKPGVTPRS 14
|:|:|:|:|
Db 20 PPIPRSGATPRS 31

RESULT 30
US-09-252-991A-17180
; Sequence 17180, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17180
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17180

Query Match 56.6%; Score 43; DB 4; Length 731;
Best Local Similarity 58.3%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 TPELPKPGVTPR 13
:|:|:|:|:|
Db 296 SPEAPRPGPRPR 307

RESULT 31
US-09-252-991A-29789
; Sequence 29789, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

```
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29789
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29789

Query Match          55.3%; Score 42; DB 4; Length 159;
Best Local Similarity 72.7%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2 TPPLPKPGVTP 12
        :|||||:
Db      2 TPPLPDPRSTP 12

RESULT 32
US-08-951-822-23
; Sequence 23, Application US/08951822A
; Patent No. 5989866
; GENERAL INFORMATION:
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Raymond, Fenella
; APPLICANT: Bukowski, Thomas R.
; APPLICANT: Holderman, Susan D.
; APPLICANT: Hansen, Birgit
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: NOVEL FGF HOMOLOGS
; FILE REFERENCE: 96-20
; CURRENT APPLICATION NUMBER: US/08/951,822A
; CURRENT FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-951-822-23

Query Match          55.3%; Score 42; DB 2; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 LPKPGVTP 12
        :|||||:
Db      218 VPKPGVTP 225

RESULT 33
US-08-705-245-4
; Sequence 4, Application US/08705245
; Patent No. 6020189
; GENERAL INFORMATION:
; APPLICANT: Nathans et al., Jeremy
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; TITLE OF INVENTION: FACTORS (FHFs) AND METHODS OF USE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/705,245
; FILING DATE: 30-AUG-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07265/094001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-50999
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-705-245-4

Query Match          55.3%; Score 42; DB 3; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 LPKPGVTP 12
        :|||||:
Db      218 VPKPGVTP 225

RESULT 34
US-09-368-951-23
; Sequence 23, Application US/09368951
; Patent No. 6352971
; GENERAL INFORMATION:
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Raymond, Fenella
; APPLICANT: Bukowski, Thomas R.
; APPLICANT: Holderman, Susan D.
; APPLICANT: Hansen, Birgit
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: NOVEL FGF HOMOLOGS
; FILE REFERENCE: 96-20
; CURRENT APPLICATION NUMBER: US/09/368,951
; CURRENT FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: 08/951,822
; EARLIER FILING DATE: 1997-10-16
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-368-951-23

Query Match          55.3%; Score 42; DB 3; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      5 LPKPGVTP 12
        :|||||:
Db      218 VPKPGVTP 225

RESULT 35
US-09-390-207-16
; Sequence 16, Application US/09390207
; Patent No. 6504530
; GENERAL INFORMATION:
; APPLICANT: Thomason, Arlen
; APPLICANT: Liu, Benxian
; TITLE OF INVENTION: Fibroblast Growth Factor-Like Polypeptides
; FILE REFERENCE: 99-371
; CURRENT APPLICATION NUMBER: US/09/390,207
```

```
; CURRENT FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-390-207-16

Query Match          55.3%; Score 42; DB 4; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      5 LPKPGVTP 12
       :|||||
Db      218 VPKPGVTP 225

RESULT 36
US-09-229-947-23
; Sequence 23, Application US/09229947
; Patent No. 6518236
; GENERAL INFORMATION:
; APPLICANT: Deisher, Theresa A.
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Raymond, Fenella
; APPLICANT: Bukowski, Thomas R.
; APPLICANT: Holderman, Susan D.
; APPLICANT: Hansen, Birgit
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: NOVEL FGF HOMOLOGS
; FILE REFERENCE: 96-20C1
; CURRENT APPLICATION NUMBER: US/09/229,947
; CURRENT FILING DATE: 1999-01-13
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-229-947-23

Query Match          55.3%; Score 42; DB 4; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      5 LPKPGVTP 12
       :|||||
Db      218 VPKPGVTP 225

RESULT 37
US-09-490-714-4
; Sequence 4, Application US/09490714
; Patent No. 6635744
; GENERAL INFORMATION:
; APPLICANT: Nathans et al., Jeremy
; TITLE OF INVENTION: FIBROBLAST GROWTH FACTOR HOMOLOGOUS
; NUMBER OF INVENTION: FACTORS (FHB) AND METHODS OF USE
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/490,714
; FILING DATE: 25-JAN-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/705,245
; FILING DATE: 30-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell, Jr., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07265/094001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-50999
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-490-714-4

Query Match          55.3%; Score 42; DB 4; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      5 LPKPGVTP 12
       :|||||
Db      218 VPKPGVTP 225

RESULT 38
US-09-949-016-9554
; Sequence 9554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9554
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9554

Query Match          55.3%; Score 42; DB 4; Length 247;
Best Local Similarity 87.5%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      5 LPKPGVTP 12
       :|||||
Db      218 VPKPGVTP 225

RESULT 39
US-08-462-169B-2
; Sequence 2, Application US/08462169B
; Patent No. 5773252
; GENERAL INFORMATION:
; APPLICANT: John Greene and Craig A. Rosen
; TITLE OF INVENTION: Fibroblast Growth Factor-15
; NUMBER OF SEQUENCES: 32
```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,169B
; FILING DATE: 05 JUN 95
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-441 (PF203)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-462-169B-2

Query Match 55.3%; Score 42; DB 1; Length 252;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 LPKPGVTP 12
:|||||
Db 223 VPKPGVTP 230

RESULT 40
US-08-462-169B-23
; Sequence 23, Application US/08462169B
; Patent No. 5773252
; GENERAL INFORMATION:
; APPLICANT: John Greene and Craig A. Rosen
; TITLE OF INVENTION: Fibroblast Growth Factor-15
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,169B
; FILING DATE: 05 JUN 95
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-441 (PF203)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700

; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-462-169B-23

Query Match 55.3%; Score 42; DB 1; Length 252;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 LPKPGVTP 12
:|||||
Db 223 VPKPGVTP 230

Search completed: October 28, 2005, 22:19:31
Job time : 39.7391 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 21:54:30 : Search time 139.391 Seconds
(without alignments)
51.432 Million cell updates/sec

Title: US-10-634-914-16
Perfect score: 76
Sequence: 1 ETEPLPKPGVTPRS 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Uniprot_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	76	100.0	488	2 Q9JS66	Q9JS66 chlamydia p
2	76	100.0	493	2 Q9Z797	Q9Z797 chlamydia p
3	55	72.4	487	2 O84582	O84582 chlamydia t
4	50	65.8	491	2 Q9PJG2	Q9PJG2 chlamydia m
5	49	64.5	59	2 Q9MX74	Q9MX74 oreochromis
6	49	64.5	59	2 Q9MX77	Q9MX77 oreochromis
7	49	64.5	59	2 Q9MX80	Q9MX80 oreochromis
8	49	64.5	62	2 Q9MX72	Q9MX72 oreochromis
9	48	63.2	59	2 Q9MX71	Q9MX71 oreochromis
10	48	63.2	59	2 Q9MX79	Q9MX79 oreochromis
11	48	63.2	65	2 Q9MX73	Q9MX73 oreochromis
12	48	63.2	65	2 Q9MX76	Q9MX76 oreochromis
13	48	63.2	65	2 Q9MX82	Q9MX82 oreochromis
14	47.5	62.5	494	2 Q821I6	Q821I6 chlamydomophi
15	47	61.8	181	2 Q82DG2	Q82DG2 streptomyc
16	47	61.8	995	2 Q9Y2W4	Q9Y2W4 homo sapien
17	47	61.8	1012	2 Q43393	Q43393 homo sapien
18	47	61.8	1012	2 Q75359	Q75359 homo sapien
19	47	61.8	1296	2 Q75046	Q75046 homo sapien
20	47	61.8	1506	2 Q6P6B9	Q6P6B9 homo sapien
21	47	61.8	1566	2 Q9P2R6	Q9P2R6 homo sapien
22	46	60.5	322	2 Q7Q062	Q7Q062 anopheles g
23	46	60.5	485	2 Q6TF31	Q6TF31 drosophila
24	46	60.5	485	2 Q9VDU9	Q9VDU9 drosophila
25	45.5	59.9	830	2 Q6CAY0	Q6CAY0 yarrowia li
26	45.5	59.9	886	2 Q7SGS4	Q7SGS4 neurospora
27	45	59.2	59	2 Q9MX70	Q9MX70 oreochromis
28	45	59.2	412	2 Q9U2Z2	Q9U2Z2 caenorhabdi
29	44	57.9	86	2 Q88NV3	Q88NV3 pseudomonas
30	44	57.9	152	2 Q7XAM1	Q7XAM1 oryza sativ
31	44	57.9	241	2 Q84TW0	Q84TW0 oryza sativ

32	44	57.9	299	2 Q96NN6	Q96NN6 homo sapien
33	44	57.9	353	2 Q6BNS6	Q6BNS6 debaryomyce
34	44	57.9	360	2 Q8BZ71	Q8BZ71 m mus muscu
35	44	57.9	482	2 Q6TF30	Q6TF30 drosophila
36	44	57.9	1257	2 P96746	P96746 corynebacte
37	44	57.9	1257	2 Q8NRC3	Q8NRC3 corynebacte
38	44	57.9	1317	2 Q7Q2L6	Q7Q2L6 anopheles g
39	44	57.9	1830	2 Q811G3	Q811G3 mus musculu
40	44	57.9	2404	1 SON_MOUSE	Q9QX47 mus musculu
41	44	57.9	2426	1 SON_HUMAN	P18583 homo sapien
42	44	57.2	419	2 Q9FHM5	Q9FHM5 arabidopsis
43	43.5	57.2	1885	2 Q8TOD1	Q8TOD1 methanobarc
44	43	56.6	130	2 Q8H599	Q8H599 oryza sativ
45	43	56.6	170	2 Q9N0P1	Q9N0P1 cercopithec
46	43	56.6	263	2 Q8H5A3	Q8H5A3 oryza sativ
47	43	56.6	330	2 Q89MF0	Q89MF0 bradyrhizob
48	43	56.6	345	2 Q51932	Q51932 peptostrept
49	43	56.6	392	1 ISPQ_MYCLE	Q9CDB5 mycobacteri
50	43	56.6	1106	2 Q6XDB6	Q6XDB6 rattus norv
51	43	56.6	1474	1 SHK2_RAT	Q9QX74 rattus norv
52	43	56.6	1476	2 Q80Z38	Q80Z38 mus musculu
53	43	56.6	1681	2 Q7BX84	Q7PX84 anopheles g
54	43	56.6	1839	2 Q6WB19	Q6WB19 rattus norv
55	43	56.6	3166	2 Q9W320	Q9W320 drosophila
56	42.5	55.9	458	2 Q7MXC3	Q7MXC3 porphyromon
57	42.5	55.9	1435	2 Q755Y6	Q755Y6 ashbya goss
58	42	55.3	111	2 Q96QX6	Q96QX6 homo sapien
59	42	55.3	174	2 Q99M83	Q99M83 mus musculu
60	42	55.3	207	2 Q7NEH2	Q7NEH2 gloebacter
61	42	55.3	224	2 Q99MF9	Q99MF9 rattus norv
62	42	55.3	247	1 FGFE_HUMAN	Q92915 homo sapien
63	42	55.3	252	2 Q86YN7	Q86YN7 homo sapien
64	42	55.3	349	2 Q91Y00	Q91Y00 mus musculu
65	42	55.3	365	2 Q6C3Y0	Q6C3Y0 yarrowia li
66	42	55.3	377	2 Q8N905	Q8N905 homo sapien
67	42	55.3	394	2 Q721L1	Q721L1 thermus the
68	42	55.3	424	2 Q6VB56	Q6VB56 human herpe
69	42	55.3	495	2 Q916C3	Q916C3 pseudomonas
70	42	55.3	533	1 VL2_HPVI5	Q916C3 human papil
71	42	55.3	535	2 Q9JL94	Q9JL94 rattus norv
72	42	55.3	548	2 Q703D2	Q703D2 human herpe
73	42	55.3	550	1 VGLE_HHV11	P04488 human herpe
74	42	55.3	550	2 Q703C1	Q703C1 human herpe
75	42	55.3	550	2 Q703C2	Q703C2 human herpe
76	42	55.3	550	2 Q703C4	Q703C4 human herpe
77	42	55.3	550	2 Q703C5	Q703C5 human herpe
78	42	55.3	550	2 Q703C7	Q703C7 human herpe
79	42	55.3	550	2 Q703C8	Q703C8 human herpe
80	42	55.3	550	2 Q703C9	Q703C9 human herpe
81	42	55.3	550	2 Q703D0	Q703D0 human herpe
82	42	55.3	550	2 Q703D1	Q703D1 human herpe
83	42	55.3	550	2 Q703D3	Q703D3 human herpe
84	42	55.3	550	2 Q703D4	Q703D4 human herpe
85	42	55.3	550	2 Q703D5	Q703D5 human herpe
86	42	55.3	550	2 Q703E0	Q703E0 human herpe
87	42	55.3	550	2 Q703E1	Q703E1 human herpe
88	42	55.3	550	2 Q703E2	Q703E2 human herpe
89	42	55.3	550	2 Q703E3	Q703E3 human herpe
90	42	55.3	550	2 Q703E5	Q703E5 human herpe
91	42	55.3	550	2 Q703E7	Q703E7 human herpe
92	42	55.3	550	2 Q703E9	Q703E9 human herpe
93	42	55.3	552	2 Q703C3	Q703C3 human herpe
94	42	55.3	552	2 Q703C6	Q703C6 human herpe
95	42	55.3	552	2 Q703D6	Q703D6 human herpe
96	42	55.3	552	2 Q703D7	Q703D7 human herpe
97	42	55.3	552	2 Q703D8	Q703D8 human herpe
98	42	55.3	552	2 Q703D9	Q703D9 human herpe
99	42	55.3	552	2 Q703E4	Q703E4 human herpe
100	42	55.3	552	2 Q703E6	Q703E6 human herpe

ALIGNMENTS

RESULT 1
Q9JS66 PRELIMINARY; PRT; 488 AA.
ID Q9JS66; Q7A193; Q7BWY2;
AC Q9JS66; Q7A193; Q7BWY2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, last annotation update)
DE Hypothetical protein CP1062 (Hypothetical protein CPj0809).
GN OrderedlocusNames=CP1062, CPJ0809, Cpb0838;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
NCBI_TaxID=83558;
OX NCBI_TaxID=83558;
RN [1]
RP STRAIN=CWL029;
RC MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Baas S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwin M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mofn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Tw-183;
RA Geng M.M., Schumacher A., Muehldorfer I., Bensch K.W., Schaefer K.P.,
RA Schneider S., Pohl T., Essig A., Marre R., Melchers K.;
RT "The genome sequence of Chlamydia pneumoniae TW183 and comparison with
RT other Chlamydia strains based on whole genome sequence analysis.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362; DOI=10.1093/nar/28.12.2311;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AE002263; AAF38835.1; -.
DR EMBL; AE017160; AAP98767.1; -.
DR EMBL; AP002548; BAA99017.1; -.
DR PIR; A81507; A81507.
DR PIR; G86591; G86591.
DR TIGR; CP1062; -.
DR InterPro; IPR000508; Peptidase S26.
DR PROSITE; PS00501; SPASE_1_1; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 488 AA; 49819 MW; 06F2F0B905398AAB CRC64;
Query Match 100.0%; Score 76; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
DB 164 ETPELPKPGVTPRS 177
RESULT 2
Q9Z797 PRELIMINARY; PRT; 493 AA.
ID Q9Z797;
AC Q9Z797;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE Hypothetical protein CPn0809.
GN OrderedlocusNames=CPn0809;
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388; DOI=10.1038/7716;
RA Kaiman S., Mitchell W.P., Marathe R., Lammel C.J., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
DR EMBL; AE001662; AAD18947.1; -.
DR PIR; F72031; F72031.
DR InterPro; IPR000508; Peptidase S26.
DR PROSITE; PS00501; SPASE_1_1; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 493 AA; 50489 MW; 90F315B7A20AC2F0 CRC64;
Query Match 100.0%; Score 76; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
DB 164 ETPELPKPGVTPRS 177
RESULT 3
O84582 PRELIMINARY; PRT; 487 AA.
ID O84582
AC O84582;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE Hypothetical protein CT578.
GN OrderedlocusNames=CT578;
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D / UW-3 / Cx;
RX MEDLINE=99000809; PubMed=9784136; DOI=10.1126/science.282.5389.754;
RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
DR EMBL; AE001328; AAC68180.1; -.
DR PIR; D71497; D71497.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 487 AA; 50217 MW; D4D6038EC1D4A03D CRC64;
Query Match 72.4%; Score 55; DB 2; Length 487;
Best Local Similarity 76.9%; Pred. No. 6.5;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPR 13
DB 161 ETPGLPKPSTTPR 173
RESULT 4
Q9PJG2 PRELIMINARY; PRT; 491 AA.
ID Q9PJG2
AC Q9PJG2;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE Hypothetical protein TC0867.
GN OrderedlocusNames=TC0867;
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPn / N199;
RX MEDLINE=20150255; PubMed=10684935; DOI=10.1093/nar/28.6.1397;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J.D., Utterback T.R., Berry K.J.,
RA Bass S., Linher K.D., Weidman J.F., Khouri H.M., Craven B., Bowman C.,
RA Dodson R.J., Gwinn M.L., Nelson W.C., DeBoy R.T., Kolonay J.F.,
RA McClarty G., Salzberg S.L., Eisen J.A., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MOPn and Chlamydia
RT pneumoniae AR39.";
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002353; AAF39663.1; -.
DR PIR; F81655; F81655.
DR TIGR; TC0867; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 491 AA; 50762 MW; 440F298930D93FB2 CRC64;

Query Match 65.8%; Score 50; DB 2; Length 491;
Best Local Similarity 69.2%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPR 13
:|||||
Db 161 DTPGLPKPTTTPR 173

RESULT 5

Q9MX74 PRELIMINARY; PRT; 59 AA.
AC Q9MX74;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Suelmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214724; AAF66864.1; -.
FT NON TER 1
SQ SEQUENCE 59 AA; 6184 MW; DB82B60BB422B35F CRC64;

Query Match 64.5%; Score 49; DB 2; Length 59;
Best Local Similarity 75.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 19 ETPETPKPSVGP 30

RESULT 6

Q9MX77 PRELIMINARY; PRT; 59 AA.
AC Q9MX77;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.

OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Suelmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214721; AAF66861.1; -.
FT NON TER 1
SQ SEQUENCE 59 AA; 6156 MW; CB59D7F28CFB340 CRC64;

Query Match 64.5%; Score 49; DB 2; Length 59;
Best Local Similarity 75.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 19 ETPETPKPSVGP 30

RESULT 7

Q9MX80 PRELIMINARY; PRT; 59 AA.
AC Q9MX80;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Suelmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214718; AAF66858.1; -.
FT NON TER 1
SQ SEQUENCE 59 AA; 6224 MW; 5C82AA24B5E46837 CRC64;

Query Match 64.5%; Score 49; DB 2; Length 59;
Best Local Similarity 75.0%; Pred. No. 5.8;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 19 ETPETPKPSVGP 30

RESULT 8

Q9MX72 PRELIMINARY; PRT; 62 AA.
AC Q9MX72;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Suelmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214726; AAF66866.1; -.
FT NON TER 1
SQ SEQUENCE 62 AA; 6525 MW; C5E6143523B49510 CRC64;

Query Match 64.5%; Score 49; DB 2; Length 62;
Best Local Similarity 75.0%; Pred. No. 6.2;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 22 ETPETPKPSVGP 33

RESULT 9

Q9MX71 PRELIMINARY; PRT; 59 AA.
AC Q9MX71;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Sueltmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214727; AAF66867.1; -.
FT NON TER 1
SQ SEQUENCE 59 AA; 6164 MW; F58C3728B864427D CRC64;

Query Match 63.2%; Score 48; DB 2; Length 59;
Best Local Similarity 75.0%; Pred. No. 8.3;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 19 ETPKPKPSVGP 30

RESULT 10

Q9MX79 PRELIMINARY; PRT; 59 AA.
AC Q9MX79;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Sueltmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214719; AAF66859.1; -.
FT NON TER 1
SQ SEQUENCE 59 AA; 6182 MW; F583B1F29011577D CRC64;

Query Match 63.2%; Score 48; DB 2; Length 59;
Best Local Similarity 75.0%; Pred. No. 8.3;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 19 ETPKPKPSVGP 30

RESULT 11
Q9MX73 PRELIMINARY; PRT; 65 AA.
AC Q9MX73;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Sueltmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214725; AAF66865.1; -.
FT NON TER 1
SQ SEQUENCE 65 AA; 6837 MW; 02F315B87EDA10C7 CRC64;

Query Match 63.2%; Score 48; DB 2; Length 65;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 25 ETPKPKPSVGP 36

RESULT 12

Q9MX76 PRELIMINARY; PRT; 65 AA.
AC Q9MX76;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Sueltmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214722; AAF66862.1; -.
FT NON TER 1
SQ SEQUENCE 65 AA; 6837 MW; 02F315B87EDA10C7 CRC64;

Query Match 63.2%; Score 48; DB 2; Length 65;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|||||
Db 25 ETPKPKPSVGP 36

RESULT 13

Q9MX82 PRELIMINARY; PRT; 65 AA.
AC Q9MX82;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;


```
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorphna; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Clariidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Sueltmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214716; AAF66856.1; -.
FT NON TER 1
SQ SEQUENCE 65 AA; 6835 MW; 4FEFD36012BE961C CRC64;

Query Match 63.2%; Score 48; DB 2; Length 65;
Best Local Similarity 75.0%; Pred. No. 9.1;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETELPKPGVTP 12
Db 25 ETEPKPKPSVGP 36

RESULT 14
Q82116 PRELIMINARY; PRT; 494 AA.
AC Q82116;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=CCA00954;
OS Chlamydomophila caviae.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydomphila.
OX NCBI_TaxID=83557;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GPIC;
RX MEDLINE=22569155; PubMed=12682364; DOI=10.1093/nar/gkg321;
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,
RA Heidelberg J.F., Holtzapfle E.K., Khouri H.M., Federova N.B.,
RA Carly H.A., Umayam L.A., Haft D.H., Peterson J.D., Beanan M.J.,
RA White O., Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G.,
RA Bavoil P.M., Fraser C.M.;
RT "Genome sequence of Chlamydomphila caviae (Chlamydia psittaci GPIC):
RT examining the role of niche-specific genes in the evolution of the
RT Chlamydiaceae."
RL Nucleic Acids Res. 31:2134-2147(2003).
DR EMBL; AE016997; AAP05693.1; -.
DR TIGR; CCA00954; -.
KW Complete proteome.
SQ SEQUENCE 494 AA; 50829 MW; C544379F027871CB CRC64;

Query Match 62.5%; Score 47.5; DB 2; Length 494;
Best Local Similarity 64.3%; Pred. No. 87;
Matches 9; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 ETELPKPGVT-PR 13
Db 163 ETPDLKPKSISKPR 176

RESULT 15
Q82DG2 PRELIMINARY; PRT; 181 AA.
AC Q82DG2;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedlocusNames=SAV5020;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
```

```
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RT Nat. Biotechnol. 21:526-531(2003).
RL EMBL; AF005041; BAC72732.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0000271; P:polysaccharide biosynthesis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007267; GtrA.
DR Pfam; PF04138; GtrA; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 181 AA; 20219 MW; 9F8887E94E3AC99B CRC64;
```

```
Query Match 61.8%; Score 47; DB 2; Length 181;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETELPKPGVTPR 13
Db 167 EEPERPRPGSSPR 179

RESULT 16
Q9Y2W4 PRELIMINARY; PRT; 995 AA.
AC Q9Y2W4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Atrophin-related protein ARP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20195630; PubMed=10729226; DOI=10.1006/geno.1999.6097;
RA Amler L.C., Bauer A., Corvi R., Dhlmann S., Praml C., Cavenee W.K.,
RA Schwab M., Hampton G.M.;
RT "Identification and characterization of novel genes located at the
RT t(1;15)(p36.2;q24) translocation breakpoint in the neuroblastoma cell
RT line NGP."
RL Genomics 64:195-202(2000).
DR EMBL; AF118275; AAD27584.1; -.
DR InterPro; IPR002951; Atrophin.
DR Pfam; PF03154; Atrophin-1; 1.
DR PRINTS; PR01222; ATROPHIN.
SQ SEQUENCE 995 AA; 107088 MW; 888B827A44BFFB04 CRC64;
```

```
Query Match 61.8%; Score 47; DB 2; Length 995;
Best Local Similarity 72.7%; Pred. No. 2.1e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TPELPKPGVTP 12
||:|||||
```

Db 187 TPQLPTPGPTP 197

RESULT 17

043393 PRELIMINARY; PRT; 1012 AA.
AC 043393;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Atrophin-1 related protein.
GN Name=DRPLA;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xia J.-H., Liu C.-Y., Ruan Q.-G., Wang D.-A., Deng H.-X.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016005; AAC31120.1; -.
DR InterPro; IPR002951; Atrophin.
DR Pfam; PF03154; Atrophin-1; 1.
DR PRINTS; PR01222; ATROPHIN.
SQ SEQUENCE 1012 AA; 108967 MW; 4D4A8F2A6B78866 CRC64;

Query Match 61.8%; Score 47; DB 2; Length 1012;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TPBLPKPGVTP 12
||:|||||
Db 204 TPQLPTPGPTP 214

RESULT 18

075359 PRELIMINARY; PRT; 1012 AA.
AC 075359;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Atrophin-1 like protein.
GN Name=ARG;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xia J.-h., Ruan Q.-g., Liu C.-y.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF041104; AAC28264.1; -.
DR EMBL; AF041096; AAC28264.1; JOINED.
DR EMBL; AF041097; AAC28264.1; JOINED.
DR EMBL; AF041098; AAC28264.1; JOINED.
DR EMBL; AF041099; AAC28264.1; JOINED.
DR EMBL; AF041100; AAC28264.1; JOINED.
DR EMBL; AF041101; AAC28264.1; JOINED.
DR EMBL; AF041102; AAC28264.1; JOINED.
DR EMBL; AF041103; AAC28264.1; JOINED.
DR InterPro; IPR002951; Atrophin.
DR Pfam; PF03154; Atrophin-1; 1.
DR PRINTS; PR01222; ATROPHIN.
SQ SEQUENCE 1012 AA; 109047 MW; 54B60DCBF0FB85FE CRC64;

Query Match 61.8%; Score 47; DB 2; Length 1012;
Best Local Similarity 72.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TPBLPKPGVTP 12
||:|||||
Db 204 TPQLPTPGPTP 214

RESULT 19

075046 PRELIMINARY; PRT; 1296 AA.
AC 075046;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE KIAA0458 protein (Fragment).
GN Name=KIAA0458;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98116662; PubMed=9455484;
RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
RA Nomura N., Ohara O.;
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
from human brain."
RL DNA Res. 4:345-349(1997).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AB007927; BAA32303.2; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0008267; F:poly-glutamine tract binding; TAS.
DR GO; GO:0006607; P:NLS-bearing substrate-nucleus import; TAS.
DR InterPro; IPR002951; Atrophin.
DR InterPro; IPR000949; ELM2.
DR InterPro; IPR001005; Myb_DNA_binding.
DR InterPro; IPR000679; Znf_GATA.
DR Pfam; PF03154; Atrophin-1; 1.
DR Pfam; PF01448; ELM2; 1.
DR Pfam; PF00320; GATA; 1.
DR Pfam; PF00249; Myb_DNA-binding; 1.
DR PRINTS; PR01222; ATROPHIN.
DR SMART; SM00717; SANT; 1.
DR SMART; SM00401; Znf_GATA; 1.
KW Nuclear protein.
FT NON_TER 1 1
SQ SEQUENCE 1296 AA; 141498 MW; 7F8D3A8009FEFC19 CRC64;

Query Match 61.8%; Score 47; DB 2; Length 1296;
Best Local Similarity 72.7%; Pred. No. 2.8e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TPBLPKPGVTP 12
||:|||||
Db 518 TPQLPTPGPTP 528

RESULT 20

06P6B9 PRELIMINARY; PRT; 1506 AA.
AC 06P6B9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE RERE protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bobak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL: BC062342; AAH62342.1; -;
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR002951; Atrophin.
DR InterPro: IPR001025; BAH.
DR InterPro: IPR000949; ELM2.
DR InterPro: IPR001005; Myb_DNA_binding.
DR InterPro: IPR000679; Znf_GATA.
DR Pfam: PF03154; Atrophin-1; 1.
DR Pfam: PF01426; BAH; 1.
DR Pfam: PF01448; ELM2; 1.
DR Pfam: PF00320; GATA; 1.
DR Pfam: PF00249; Myb_DNA_binding; 1.
DR PRINTS: PR01222; ATROPHIN.
DR SMART: SM00439; BAH; 1.
DR SMART: SM00717; SANT; 1.
DR SMART: SM00401; Znf_GATA; 1.
KW Nuclear protein.
SQ SEQUENCE 1506 AA; 165609 MW; 3BE4D9C2BE8772AA CRC64;

Query Match 61.8%; Score 47; DB 2; Length 1506;
Best Local Similarity 72.7%; Pred. No. 3.3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TPPLPKPGVTP 12
Db 698 TPQLPTPGPTP 708

RESULT 21
Q9P2R6 PRELIMINARY; PRT; 1566 AA.
ID Q9P2R6;
AC Q9P2R6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE RERE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20275632; Pubmed=10814707; DOI=10.1093/hmg/9.9.1433;
RA Yanagisawa H., Bundo M., Miyashita T., Okamura-Oho Y., Tadokoro K.,
RA Tokunaga K., Yamada M.;
RT "Protein binding of a DRPLA family through arginine-glutamic acid
dipeptide repeats is enhanced by extended polyglutamine.";
RL Hum. Mol. Genet. 9:1433-1442(2000).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).

DR EMBL: AB036737; BAA95898.1; -;
DR Genew; HGNC:9965; RERE.
DR GO: GO:0005634; C:nucleus; NAS.
DR GO: GO:0005515; F:protein binding; NAS.
DR InterPro: IPR002951; Atrophin.
DR InterPro: IPR001025; BAH.
DR InterPro: IPR000949; ELM2.
DR InterPro: IPR001005; Myb_DNA_binding.
DR InterPro: IPR000679; Znf_GATA.
DR Pfam: PF03154; Atrophin-1; 1.
DR Pfam: PF01426; BAH; 1.
DR Pfam: PF01448; ELM2; 1.
DR Pfam: PF00320; GATA; 1.
DR Pfam: PF00249; Myb_DNA_binding; 1.
DR SMART: SM00439; BAH; 1.
DR SMART: SM00717; SANT; 1.
DR SMART: SM00401; Znf_GATA; 1.
KW Nuclear protein.
SQ SEQUENCE 1566 AA; 172336 MW; 4047F1530F198E2C CRC64;

Query Match 61.8%; Score 47; DB 2; Length 1566;
Best Local Similarity 72.7%; Pred. No. 3.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TPPLPKPGVTP 12
Db 758 TPQLPTPGPTP 768

RESULT 22
Q7Q062 PRELIMINARY; PRT; 322 AA.
ID Q7Q062;
AC Q7Q062;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE AgCP9323 (Fragment).
GN Name=agCG55833; ORFNames=ENSANGG00000014181;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 3 WD repeats.
CC -1- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
preliminary data.
CC EMBL: AAB01008986; EAA00573.1; -;
DR EMBL: AAB01008986; EAA00573.1; -;
DR InterPro: IPR001680; WD40.
DR InterPro: IPR011046; WD40_like.
DR Pfam: PF00400; WD40; 3.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
KW Repeat; WD repeat.
FT NON TER 322
SQ SEQUENCE 322 AA; 35311 MW; C1BB5AF78CBB9E14 CRC64;

Query Match 60.5%; Score 46; DB 2; Length 322;
Best Local Similarity 69.2%; Pred. No. 34;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ETPLPKPGVTPR 13
Db 148 ETPEDTKPGTTPR 160

RESULT 23
Q6TF31 PRELIMINARY; PRT; 485 AA.
ID Q6TF31;
AC Q6TF31;

DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Histamine-gated chloride channel.
GN Name=hclA;
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7240;
RN [1]
RP SEQUENCE FROM N.A.
RA Iovchev M., Boutanaev A., Wolstenholme A., Nurminsky D., Semenov E.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the ligand-gated ionic channel (TC 1.A.9) family.
CC
DR EMBL; AY422812; AAR33080.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045211; C:postsynaptic membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0004890; F:GABA-A receptor activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006028; GABA_recept.
DR InterPro; IPR006201; Neur_channel.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00253; GABARECEPT.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; UNKNOWN 1.
KW Ion transport; Ionic channel; Postsynaptic membrane; Transmembrane;
KW Transport.
SQ SEQUENCE 485 AA; 55550 MW; 62DFD3849E3747C5 CRC64;
QY 3 PELPKPGVTP 12
Db 438 PELPKPKLTP 447
Query Match 60.5%; Score 46; DB 2; Length 485;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
RESULT 24
ID Q9VDU9 PRELIMINARY; PRT; 485 AA.
AC Q9VDU9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE CG7411-PA (Histamine-gated chloride channel subunit 2) (Histamine-gated chloride channel subunit 1) (Histamine-gated chloride channel subunit A).
DE gated chloride channel subunit 1)
GN Name=ort; Synonyms=Hist1, hclA; ORFNames=CG7411;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazewicz R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle E.G., Baxter B.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodagel, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleab J.M., Park S., Pfeiffer B.D., Richards S., Sodegren E.J.,
RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
RA Smith C.D., Tupu J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Beltencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.


```
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=21648756; PubMed=11714703; DOI=10.1074/jbc.M107635200;
RA Zheng Y., Hirschberg B., Yuan J., Wang A.P., Hunt D.C., Ludmerer S.W.,
RA Schmatz D.M., Cully D.F.;
RT "Identification of two novel Drosophila melanogaster histamine-gated
RT chloride channel subunits expressed in the eye.";
RL J. Biol. Chem. 277:2000-2005(2002).
RN [8]
RP SEQUENCE FROM N.A.
RA Witte I., Kreienkamp H.-J., Gewecke M., Roeder T.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=21624647; PubMed=11753412; DOI=10.1038/n787;
RA Gisselmann G., Pusch H., Hovemann B.T., Hatt H.;
RT "Two cDNAs coding for histamine-gated ion channels in D.
RT melanogaster.";
RL Nat. Neurosci. 5:11-12(2002).
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon R;
RA Yin Z., Skingsley D.R., Geng C., Bowman J., Koliantz G., Burg M.G.,
RA Semenov E.P., Hardie R.C., Pak W.L.;
RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the ligand-gated ionic channel (TC 1.A.9)
CC family.
DR EMBL; AE003727; AAF55691.1; -
DR EMBL; AF382403; AAL66188.1; -
DR EMBL; AF411340; AAL05873.1; -
DR EMBL; AF435469; AAL74413.1; -
DR EMBL; AY049774; AAL12210.1; -
DR HSSP; P23415; IMOT.
DR FlyBase; FBgn003011; ort.
DR GO; GO:0030534; P:adult behavior; IMP.
DR GO; GO:0045472; P:response to ether; IMP.
DR GO; GO:0009636; P:response to toxin; IMP.
DR InterPro; IPR006028; GABAA_recept.
DR InterPro; IPR006201; Neur_channel.
DR InterPro; IPR006202; Neur_channel_LBD.
DR InterPro; IPR006029; Neu_channel_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00253; GABAARCEPTR.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; UNKNOWN 1.
KW Ion transport; Ionic channel; Postsynaptic membrane; Transmembrane;
KW Transport.
SQ SEQUENCE 485 AA; 55547 MW; 291ECE295C5E114E CRC64;

Query Match 60.5%; Score 46; DB 2; Length 485;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
Db 438 PELPKPKLTP 447

RESULT 25
O6CAY0 PRELIMINARY; PRT; 830 AA.
AC O6CAY0;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Similarity.
GN ORFNames=YALI0C23452g;
OS Yarrowia lipolytica CLIB99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
```

```
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.;
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Catolico L., Confantolieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hanttraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIB99;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382129; CAG82502.1; -
DR InterPro; IPR004829; Csurface_antigen.
DR ProDom; PD153432; Csurface_antigen; 2.
SQ SEQUENCE 830 AA; 85939 MW; E0A2CF14196AF404 CRC64;
```

```
Query Match 59.9%; Score 45.5; DB 2; Length 830;
Best Local Similarity 76.9%; Pred. No. 3e+02;
Matches 10; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
```

```
QY 1 ETEPLPKP-GVTP 12
Db 596 ETEPTPKPTGETP 608
```

```
RESULT 26
Q7SGS4
ID Q7SGS4 PRELIMINARY; PRT; 886 AA.
AC Q7SGS4;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU08338.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehnman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysseis M., Mauceli E., Bieleke C., Rudd S., Frisman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
```

CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL; AABX0100008; EAA36005.1; -.
DR InterPro; IPR009058; wing_hlx_DNA_bnd.
SQ SEQUENCE 886 AA; 100251 MW; 19F4ADEFC5B3CD5E CRC64;

Query Match 59.9%; Score 45.5; DB 2; Length 886;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

OY 2 TPPLPKP-GVTPRS 14
Db 396 TPVPPKPKGDVPRS 409

RESULT 27

O9MX70 PRELIMINARY; PRT; 59 AA.

AC O9MX70; 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II alpha chain (Fragment).
GN Name=Orni-DCA;
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;

RN [1]
RP SEQUENCE FROM N.A.
RA Murray B.W., Shintani S., Suelmann H., Klein J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF214728; AAF66868.1; -.
FT NON TER 1 1
SQ SEQUENCE 59 AA; 6168 MW; CB59D964C32FB340 CRC64;

Query Match 59.2%; Score 45; DB 2; Length 59;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ETPPLPKPGVTP 12
Db 19 ETPETPPQPSVGP 30

RESULT 28

O9U2Z2 PRELIMINARY; PRT; 412 AA.

AC O9U2Z2; 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein Y113G7A.11.
GN Name=ssu-1; ORFNames=Y113G7A.11;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology."
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Lennard N.;

RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL132858; CAB60475.1; -.
DR HSSP; Q06520; 1EFH.
DR WormBase; WBGene00013748; ssu-1.
DR WormPep; Y113G7A.11; CE23282.
DR GO; GO:0008146; F:sulfoltransferase activity; IEA.
DR InterPro; IPR000863; Sulfoltransferase.
DR Pfam; PF00685; Sulfoltransfer_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 412 AA; 47251 MW; 2D3062EDB9911DC4 CRC64;

Query Match 59.2%; Score 45; DB 2; Length 412;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ETPPLPKPGVTPR 13
Db 4 KTKPKPKPPQTPR 16

RESULT 29

O88NV3 PRELIMINARY; PRT; 86 AA.

AC O88NV3; 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PP1101;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzaple E.K., Scanlan D., Tran K.,
RA Wozarz A., Utterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hohnsels J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016778; AAN66726.1; -.
DR TIGR; PP1101; -.
KW Complete proteome; Hypothetical protein.

Query Match 57.9%; Score 44; DB 2; Length 86;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPPLPKPGVTP 12
Db 56 DTPQPEPGETP 67

RESULT 30

O7XAM1 PRELIMINARY; PRT; 152 AA.

AC O7XAM1; 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Proline-rich protein family-like protein.
GN Name=OJ1634_B10.111;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ1634 B10."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003840; BAC81167.1; -.
DR Gramene; Q7XAM1; -.
SQ SEQUENCE 152 AA; 16298 MW; 70B5593BFD7FBBF9 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 152;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
Db 49 ELPELPKPELP 60

RESULT 31
Q84TW0 PRELIMINARY; PRT; 241 AA.
AC Q84TW0;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Putative AP2 domain containing protein (Putative AP2 domain
DE transcripction factor).
GN Name=OSJNBa0094J08.32; Synonyms=OSJNB0042K11.4;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tseltrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tseltrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC133007; AAO60030.1; -.
DR EMBL; AC139172; AAR01751.1; -.
DR HSSP; O80337; 2GCC.
DR Gramene; Q84TW0; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001471; TF_BRF.
DR PRINTS; PR00367; ETHRSPELEMT.
DR ProDom; PD001423; TF_ERF; 1.
```

```
SQ SEQUENCE 241 AA; 25487 MW; CAC33EA65515EAB3 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 241;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
Db 147 PELPKPTISP 156

RESULT 32
Q96NN6 PRELIMINARY; PRT; 299 AA.
AC Q96NN6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein FLJ30483.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hoshiro T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fukuiwa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itch T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK055045; BAB70844.1; -.
DR HSSP; P20480; 1N6M.
DR GO; GO:0005875; C:microtubule associated complex; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR InterPro; IPR002453; Beta_tubulin.
DR InterPro; IPR001752; kinesin_motor.
DR InterPro; IPR010989; t-snare.
DR Pfam; PF00225; kinesin; 1.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS50067; KINESIN MOTOR DOMAIN2; 1.
DR PROSITE; PS00228; TUBULIN_B AUTOREG; UNKNOWN_1.
KW ATP-binding; Microtubule; Motor protein.
SQ SEQUENCE 299 AA; 32600 MW; 4893497BESD7E08C CRC64;
```


Query Match 57.9%; Score 44; DB 2; Length 299;
Best Local Similarity 58.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETELPKPGVTP 12
|:|:|:|:|:
Db 281 ESPEHPRFGLPP 292

RESULT 33

O6BNS6 PRELIMINARY; PRT; 353 AA.

AC O6BNS6; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Similar to sp|P17423 Saccharomyces cerevisiae YHR025w THRI homoserine
kinase.
GN ORFNames=DEHA0E204059;
OS Debaryomyces hanseni CBS767.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX NCBI_TaxID=284592;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genolevures;

RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
LaFontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
Boisrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
Hantreya F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
Nicaud J.M., Nikolski M., Oztas S., Ozler-Kalogeropoulos O.,
Pelienz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
Wincker P., Souciet J.L.;
RA "Genome evolution in yeasts."
RT Nature 430:35-44(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS767;
RG Genoscope;

RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.

RL EMBL; CR382137; CAG88417.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004413; F:homoserine kinase activity; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0016310; P:phosphorylation; IEA.
DR GO; GO:0006566; P:threonine metabolism; IEA.
DR InterPro; IPR006203; GHMPkinse_ATP.
DR InterPro; IPR006204; GHMP_kinase.
DR InterPro; IPR000870; Homoser_kin.
DR Pfam; PF00288; GHMP_kinases_1.
DR PIRSF; PIRSF000676; Homoser_kin; 1.
DR PRINTS; PR00958; HOMSERKINASE.
DR TIGRFAMs; TIGR00191; thrB; 1.
DR PROSITE; PS00627; GHMP_KINASES_ATP; UNKNOWN_1.
KM Kinase.
SQ SEQUENCE 353 AA; 38822 MW; E996028AE1PA6ACE CRC64;

Query Match 57.9%; Score 44; DB 2; Length 353;
Best Local Similarity 88.9%; Pred. No. 2.1e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 LPKPGVTPR 13
|||||
Db 179 LPKPGKTPR 187

RESULT 34
O8BZ71 PRELIMINARY; PRT; 360 AA.

AC O8BZ71; 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Mus musculus adult male bone cDNA, RIKEN full-length enriched library,
clone:9830125E18 product:CDNA FLJ32451 FIS, CLONE SKMUS2001668, WEAKLY
DE SIMILAR TO NEURON-SPECIFIC SIGNAL TRANSDUCTION PROTEIN STAC homolog
DE (9830125E18 protein).
GN Name=9830125E18;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RA the FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Konno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
Nihi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyra T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Jaw and limb;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE FROM N.A.
RC TISSUE=Jaw and limb;
RA Strausberg R.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 SH3 domain.
DR EMBL; AK036516; BAC29460.1; -.
DR EMBL; BC067208; AAH67208.1; -.
DR HSSP; O89100; IOEB.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001018; Neu_cyt_fact_2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR011511; SH3_2.
DR Pfam; PF00130; C1_1; 1.
DR Pfam; PF00018; SH3_1; 1.
DR Pfam; PF07653; SH3_2; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR PRINTS; PR00499; P67PHOX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00326; SH3; 2.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50002; SH3; 1.
KW SH3 domain.
SQ SEQUENCE 360 AA; 41000 MW; ADA948B421F5B0F8 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 360;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DE Histamine-gated chloride channel.
GN Name=hclA;
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RA Iovchev M., Boutanaev A., Wolstenholme A., Nurminsky D., Semenov E.;
RT "Drosophila virilis histamine-gated chloride channel gene (hclA-DV),
RT the exon-intron sequence of ORF.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel (TC 1.A.9)
CC family.
DR EMBL; AY422813; AAR33081.2; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0045211; C:postsynaptic membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0004890; F:GABA-A receptor activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
KW Ion transport; Ionic channel; Postsynaptic membrane; Transmembrane;
KW Transport.
SQ SEQUENCE 482 AA; 55204 MW; BC90BFFBD6D0695D CRC64;

Query Match 57.9%; Score 44; DB 2; Length 482;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
Db 435 PEMPKPKLTP 444

RESULT 36
P96746
ID P96746 PRELIMINARY; PRT; 1257 AA.
AC P96746;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE 2-oxoglutarate dehydrogenase (EC 1.2.4.2).
GN Name=odhA;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterinae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AJ12036;
RX MEDLINE=97158227; PubMed=9004499;
RA Usuda Y., Tujimoto N., Abe C., Asakura Y., Kimura E., Kawahara Y.,
RA Kurahashi O., Matsui H.;
RT "Molecular cloning of the Corynebacterium glutamicum ('Brevibacterium
RT lactofermentum' AJ12036) odhA gene encoding a novel type of 2-
RT oxoglutarate dehydrogenase.";
RL Microbiology 142:3347-3354(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AJ12036;
RA Usuda Y., Tujimoto N., Abe C., Kimura E., Kawahara Y., Kurahashi O.,
RA Matsui H.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
DR EMBL; D84102; BAA12222.1; -.
DR HSSP; P07016; 1E2O.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0004591; F:oxoglutarate dehydrogenase (succinyl-transf. . .; IEA.

DR GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2Oxoacid_dh.
DR InterPro; IPR001017; Dehydrogenase_El.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR005475; Transketolase_CR.
DR Pfam; PF00198; 2-oxoacid_dh; 1.
DR Pfam; PF00676; El_dh; 1.
DR Pfam; PF02779; Transket_pyr; 1.
DR ProDom; PD001115; 2Oxoacid_dh; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 1257 AA; 138657 MW; BDE827E8B0EFB160 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 1257;
Best Local Similarity 58.3%; Pred. No. 7.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
:|:|:|:|:|
Db 131 QQPKLPEPGQTP 142

RESULT 37
Q8NRC3 PRELIMINARY; PRT; 1257 AA.
ID Q8NRC3 Q6M641;
AC Q8NRC3; Q6M641;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Pyruvate and 2-oxoglutarate dehydrogenases, El component (EC 1.2.4.2)
DE (2-OXOGLUTARATE DEHYDROGENASE).
GN Name=odha; OrderedlocusNames=Cg11129, cg1280;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahm H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins.";
RL J. Biotechnol. 104:5-25(2003).
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
DR EMBL; AP005277; BAB98522.1; -.
DR EMBL; BX927151; CAF19835.1; -.
DR HSSP; P07016; IE20.
DR GO:0008415; F:acyltransferase activity; IEA.
DR GO:0016491; F:oxidoreductase activity; IEA.
DR GO:0004591; F:oxoglutarate dehydrogenase activity; IEA.
DR GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001078; 2Oxoacid_dh.
DR InterPro; IPR001017; Dehydrogenase_El.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR005475; Transketolase_CR.
DR Pfam; PF00198; 2-oxoacid_dh; 1.
DR Pfam; PF00676; El_dh; 1.
DR Pfam; PF02779; Transket_pyr; 1.
DR ProDom; PD001115; 2Oxoacid_dh; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

KW Complete proteome; Oxidoreductase; Pyruvate.
SQ SEQUENCE 1257 AA; 138756 MW; 96473BCB12B22ABF CRC64;

Query Match 57.9%; Score 44; DB 2; Length 1257;
Best Local Similarity 58.3%; Pred. No. 7.6e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
:|:|:|:|:|
Db 131 QQPKLPEPGQTP 142

RESULT 38
Q7Q2L6 PRELIMINARY; PRT; 1317 AA.
ID Q7Q2L6 Q7Q2L6;
AC Q7Q2L6;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Ebip3627 (Fragment).
GN Name=ebig3627; ORFNames=ENSANGG00000002905;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008968; EAA13296.1; -.
DR HSSP; Q9WV48; 1Q30.
DR GO:0005515; F:protein binding; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CalB.
DR InterPro; IPR001478; PDZ.
DR InterPro; IPR000219; RhogEF.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00595; PDZ; 1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50010; DH_2; 1.
DR PROSITE; PS50106; PDZ; 1.
FT NON TER 1
FT NON TER 1
FT NON TER 1317
SQ SEQUENCE 1317 AA; 146724 MW; 9754561F35A3DC64 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 1317;
Best Local Similarity 61.5%; Pred. No. 8e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPPELPKPGVTPRS 14
:|:|:|:|:|
Db 1213 TPETPPNVSPSS 1225

RESULT 39
Q81IG3 PRELIMINARY; PRT; 1830 AA.
ID Q81IG3 Q81IG3;
AC Q81IG3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Son protein (Fragment).
GN Name=Son;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dlatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Skimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Director MGC Project;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046419; AAH46419.1;
DR GO; GO:0005515; F:protein binding; IPI.
FT NON_TER 1830 1830
SQ SEQUENCE 1830 AA; 195285 MW; B55A4992F1DCEDCE CRC64;

QY 3 PEPPKPGVTP 12
Db 431 PELFGPSVTP 440

Query Match 57.9%; Score 44; DB 2; Length 1830;
Best Local Similarity 80.0%; Pred No. 1.1e+03;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 40
SON_MOUSE STANDARD; PRT; 2404 AA.
ID SON_MOUSE
AC Q9QX47; Q9QC12; Q9COK6; Q9QXP5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE SON protein.
GN Name=Son;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=129/SV;
RX MEDLINE=20408886; PubMed=10950926; DOI=10.1006/geno.2000.6254;
RA Wynn S.L., Fisher R.A., Pagel C., Price M., Liu O.Y., Khan I.M.,
RA Zammitt P., Dadrah K., Mazrani W., Kessling A., Lee J.S., Bulwela L.,
RT "Organization and conservation of the GART/SON/DONSON locus in mouse
RT and human genomes.";
RL Genomics 68:57-62(2000).
RN [2]
RP SEQUENCE OF 1-116 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hippocampus, Small intestine, and Tongue;
RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matcoda H., Batalov S., Beisel K.W.,

RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
CC -!- FUNCTION: Transcriptional repressor. Binds to the consensus DNA
CC sequence: 5'-GA[GT]AN[CG]AG[CC]-3'. Might protect cells from
CC apoptosis. Might be involved in pre-mRNA splicing (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9QX47-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9QX47-2; Sequence=VSP_004416, VSP_004417;
CC -!- TISSUE SPECIFICITY: Widely expressed.
CC -!- DOMAIN: Contains 8 types of repeats which are distributed in 3
CC regions.
CC -!- SIMILARITY: Contains 1 DRBM (double-stranded RNA-binding) domain.
CC -!- SIMILARITY: Contains 1 G-patch domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF193606; AAF23120.1; -
CC EMBL; AF193595; AAF23120.1; JOINED.
CC EMBL; AF193596; AAF23120.1; JOINED.
CC EMBL; AF193597; AAF23120.1; JOINED.
CC EMBL; AF193598; AAF23120.1; JOINED.
CC EMBL; AF193599; AAF23120.1; JOINED.
CC EMBL; AF193600; AAF23120.1; JOINED.
CC EMBL; AF193601; AAF23120.1; JOINED.
CC EMBL; AF193602; AAF23120.1; JOINED.
CC EMBL; AF193603; AAF23120.1; JOINED.
CC EMBL; AF193604; AAF23120.1; JOINED.
CC EMBL; AF193605; AAF23120.1; JOINED.
CC EMBL; AF193607; AAF23121.1; -
CC EMBL; AK019312; BAB31659.1; -
CC EMBL; AK019081; BAB31536.1; -
CC EMBL; AK008478; BAB25691.1; -
CC EMBL; AK008256; BAB25562.1; -
CC MGD; MGI:98353; Son.
CC GO; GO:0005515; F:protein binding; IPI.
CC InterPro; IPR001159; DS_RBD.
CC InterPro; IPR000467; G_patch.
CC Pfam; PF00035; dsrm; 1.
CC Pfam; PF01585; G_patch; 1.
CC SMART; SM00443; G_patch; 1.
CC PROSITE; PS0137; DS_RBD; 1.

DR	PROSITE; PS50174; G_PATCH; 1.	
KW	Alternative splicing; DNA-binding; Nuclear protein; Repeat;	
KW	RNA-binding.	
FT	DOMAIN 721 850	13 X 10 AA tandem repeats of L-A-[ST]-[NSG]-[TS]-MDSQM.
FT	DOMAIN 867 943	11 X 7 AA tandem repeats of [DR]-P-Y-R-[LI]-[AG]-[QHP].
FT	DOMAIN 961 1080	14 X 6 AA repeats of [ED]-R-S-M-M-S.
FT	REPEAT 961 966	1-1.
FT	REPEAT 969 974	1-2.
FT	REPEAT 976 981	1-3.
FT	REPEAT 985 990	1-4.
FT	REPEAT 993 998	1-5.
FT	REPEAT 1001 1006	1-6.
FT	REPEAT 1010 1015	1-7.
FT	REPEAT 1018 1023	1-8.
FT	REPEAT 1026 1031	1-9.
FT	REPEAT 1035 1040	1-10.
FT	REPEAT 1044 1049	1-11.
FT	REPEAT 1055 1060	1-12.
FT	REPEAT 1066 1071	1-13.
FT	REPEAT 1075 1080	1-14.
FT	DOMAIN 1101 1133	3 X 11 AA tandem repeats of P-P-L-P-P-E-E-P-P-[TME]-[MTG].
FT	DOMAIN 1910 1979	7 X 7 AA repeats of P-S-R-R-S-R-[TS].
FT	REPEAT 1910 1916	2-1.
FT	REPEAT 1938 1944	2-2.
FT	REPEAT 1945 1951	2-3.
FT	REPEAT 1952 1958	2-4.
FT	REPEAT 1959 1965	2-5.
FT	REPEAT 1966 1972	2-6.
FT	REPEAT 1973 1979	2-7 (approximate).
FT	DOMAIN 1919 1990	2 X 19 AA repeats of P-S-R-R-R-S-R-S-V-V-R-R-R-S-F-S-I-S.
FT	REPEAT 1919 1937	3-1.
FT	REPEAT 1980 1990	3-2 (approximate).
FT	DOMAIN 1991 2017	3 X tandem repeats of [ST]-P-[VLI]-R-[RL]-[RK]-[RF]-S-R-G-patch.
FT	DOMAIN 2283 2329	DRBM.
FT	DOMAIN 2349 2404	K -> F (in isoform 2).
FT	VARSPPLIC 2086 2086	/FTId=VSP_004416.
FT	VARSPPLIC 2087 2404	Missing (in isoform 2).
FT	VARSPPLIC 2087 2404	/FTId=VSP_004417.
FT	SEQUENCE 2404 AA; 261428 MW; 648BF28ED3FC01D9 CRC64;	
QY	3 PELPKPGVTP 12	Query Match 57.9%; Score 44; DB 1; Length 2404;
Db	431 PELPGPSVTP 440	Best Local Similarity 80.0%; Pred. No. 1.5e+03;
		Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Search completed: October 28, 2005, 22:15:47
Job time : 145.391 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 22:02:57 ; Search time 28.6087 Seconds
(without alignments)
47.085 Million cell updates/sec

Title: US-10-634-914-16
Perfect score: 76
Sequence: 1 ETPELPKPGVTPRS 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	76	100.0	488	2	AB1507	conserved hypothet
2	76	100.0	488	2	G86591	CT578 hypothetical
3	76	100.0	493	2	F72031	ct578 hypothetical
4	55	72.4	487	2	D71497	hypothetical prote
5	50	65.8	491	2	F81655	conserved hypothet
6	43	56.6	345	2	S55377	urPAB protein prec
7	43	56.6	392	2	G87106	conserved hypothet
8	43	56.6	1252	2	T14272	cortactin-binding
9	42	55.3	495	2	C83598	hypothetical prote
10	42	55.3	533	2	S36477	l2 protein - human
11	42	55.3	550	1	VGBE18	glycoprotein E - h
12	42	55.3	603	2	T34901	probable gamma-glu
13	42	55.3	1164	1	FCSOAG	IgA Fc receptor pr
14	42	55.3	1479	2	T42710	mannose receptor,
15	42	55.3	1664	2	T18262	S-layer protein -
16	41	53.9	216	2	S73818	hypothetical prote
17	41	53.9	519	2	H86378	protein F21J9.16 l
18	41	53.9	530	2	B87369	long-chain-fatty-a
19	41	53.9	1033	2	F81595	hypothetical prote
20	41	53.9	1043	2	G86550	hypothetical prote
21	41	53.9	1043	2	G72073	hypothetical prote
22	41	53.9	1219	2	T06608	disease resistance
23	40	52.6	152	2	S76971	hypothetical prote
24	40	52.6	224	2	C48652	transfer protein s
25	40	52.6	261	2	T46233	hypothetical prote
26	40	52.6	275	2	D85070	hypothetical prote
27	40	52.6	416	2	A45510	probable protein k
28	40	52.6	495	2	T28717	hypothetical prote
29	40	52.6	644	2	A25684	hypothetical prote

30	40	52.6	728	2	S43768	transcription acti
31	40	52.6	886	1	A47521	capsid protein - g
32	40	52.6	963	2	T19140	hypothetical prote
33	40	52.6	1870	2	C47521	gag-pol-like fusio
34	40	52.6	2471	2	T42977	large tegument pro
35	40	52.6	4957	2	T03455	AlR protein - huma
36	39	51.3	144	2	T12738	ATP/GTP-binding pr
37	39	51.3	145	2	T17931	IgA Fc receptor-li
38	39	51.3	172	2	B26414	95K nonspecific cr
39	39	51.3	177	1	C40428	nonspecific cross-
40	39	51.3	182	2	F83453	adenine phosphorib
41	39	51.3	241	1	JQ1934	GTP cyclolhydrolase
42	39	51.3	321	2	JH0395	biliary glycoprote
43	39	51.3	341	2	AC0579	(citrate (pro-3S) -
44	39	51.3	344	2	A27681	nonspecific cross-
45	39	51.3	351	2	JH0396	biliary glycoprote
46	39	51.3	368	2	H69335	iron-sulfur cluste
47	39	51.3	379	2	T19069	hypothetical prote
48	39	51.3	388	2	JC5437	spliceosome-associ
49	39	51.3	417	2	JH0394	biliary glycoprote
50	39	51.3	464	2	C30127	transmembrane carc
51	39	51.3	478	2	C29514	muscarinic acetylch
52	39	51.3	479	2	S33776	muscarinic acetylch
53	39	51.3	526	1	A32164	biliary glycoprote
54	39	51.3	553	2	T45872	hypothetical prote
55	39	51.3	613	2	A56031	potassium channel
56	39	51.3	656	2	A41870	dnaA protein - Str
57	39	51.3	702	2	A36319	carcinoembryonic a
58	39	51.3	736	2	I51691	dishevelled homolo
59	39	51.3	862	2	T34342	hypothetical prote
60	39	51.3	865	2	A47282	calcium-binding pr
61	39	51.3	873	2	A47283	calphorin - fruit
62	39	51.3	950	2	F86286	hypothetical prote
63	39	51.3	1006	2	T42731	atrophin-1 related
64	39	51.3	1085	2	S66149	gene pipsqueak pro
65	39	51.3	1615	2	JC6510	ras-responsive ele
66	38.5	50.7	182	2	S75533	bidirectional hydr
67	38.5	50.7	184	2	S74232	hypothetical prote
68	38.5	50.7	499	2	A12449	glycoprotein B pre
69	38.5	50.7	917	1	VGEBBH	Ig heavy chain V r
70	38.5	50.7	1053	2	T30937	hypothetical prote
71	38	50.0	104	2	PH0991	probable proline-r
72	38	50.0	111	2	B32476	hypothetical prote
73	38	50.0	134	2	D84887	probable regulator
74	38	50.0	136	2	T35335	hypothetical prote
75	38	50.0	162	2	AB0459	adenine phosphorib
76	38	50.0	168	2	S64830	adenine phosphorib
77	38	50.0	172	2	S75440	adenine phosphorib
78	38	50.0	172	2	AF2378	muramidase (import
79	38	50.0	233	2	D86603	muramidase (invasi
80	38	50.0	233	2	A72022	methylenetetrahydr
81	38	50.0	283	2	E69626	hypothetical prote
82	38	50.0	294	2	A12016	ISA0963-2 transpos
83	38	50.0	299	2	E69288	beta-lactamase - p
84	38	50.0	313	2	A48903	hypothetical prote
85	38	50.0	344	2	AH1904	serine/threonine-s
86	38	50.0	380	2	S32831	acetyl-CoA C-acety
87	38	50.0	394	2	B48376	hypothetical prote
88	38	50.0	412	2	D86203	probable low calci
89	38	50.0	421	2	A71558	MTM1 protein - yea
90	38	50.0	433	2	S46668	hypothetical prote
91	38	50.0	446	2	G85064	acetyl-CoA carboxy
92	38	50.0	449	1	JS0632	hypothetical prote
93	38	50.0	449	2	D85990	biotin carboxylase
94	38	50.0	449	2	A10912	hypothetical prote
95	38	50.0	449	2	H91144	zinc finger protei
96	38	50.0	460	2	JE0367	glycoprotein C - h
97	38	50.0	511	1	VGBE1K	hypothetical prote
98	38	50.0	511	1	VGBEF4	probable secreted
99	38	50.0	515	2	F70904	
100	38	50.0	519	2	B87233	

ALIGNMENTS

```
RESULT 1
A81507
conserved hypothetical protein CP1062 [imported] - Chlamydophila pneumoniae (strain AR39
C:/Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:/Accession: A81507
R;/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberger, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;/Reference number: A81500; MUID:20150255; PMID:10684935
A;/Accession: A81507
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-488 <REA>
A;/Cross-references: UNIPROT:Q9JS66; GB:AE002263; GB:AE002161; NID:g7189971; PIDN:AAF3883
A;/Experimental source: strain AR39, HL cells
C:/Genetics:
A;/Gene: CP1062
C;/Superfamily: conserved hypothetical protein TC0867

Query Match          100.0%; Score 76; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||||||
Db       164 ETPELPKPGVTPRS 177

RESULT 2
G86591
CT578 hypothetical protein [imported] - Chlamydophila pneumoniae (strain J138)
C:/Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:/Accession: G86591
R;/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A;/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;/Reference number: A86491; MUID:20330349; PMID:10871362
A;/Accession: G86591
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-488 <STO>
A;/Cross-references: UNIPROT:Q9JS66; GB:BA000008; NID:g8979183; PIDN:BA99017.1; GSPDB:GN
A;/Experimental source: strain J138
C:/Genetics:
A;/Gene: CPJ0809
C;/Superfamily: conserved hypothetical protein TC0867

Query Match          100.0%; Score 76; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.00061;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||||||
Db       164 ETPELPKPGVTPRS 177

RESULT 3
F72031
ct578 hypothetical protein - Chlamydophila pneumoniae (strain CWL029)
C:/Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C:/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:/Accession: F72031
R;/Kalman, S.; Mitchell, W.; Marache, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A;/Reference number: A72000; MUID:99206606; PMID:10192388
A;/Accession: F72031
```

```
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-493 <ARN>
A;/Cross-references: UNIPROT:Q9Z797; GB:AE001662; GB:AE001363; NID:g4377118; PIDN:AAD189
A;/Experimental source: strain CWL029
C;/Genetics:
A;/Gene: CPn0809
C;/Superfamily: conserved hypothetical protein TC0867

Query Match          100.0%; Score 76; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||||||
Db       164 ETPELPKPGVTPRS 177

RESULT 4
D71497
hypothetical protein CT578 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:/Species: Chlamydia trachomatis
C:/Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:/Accession: D71497
R;/Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marache, R.; Aravind, L.; Mitchell
Science 282, 754-759, 1998
A;/Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra
A;/Reference number: A71570; MUID:9900809; PMID:9784136
A;/Accession: D71497
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-487 <ARN>
A;/Cross-references: UNIPROT:O84582; GB:AE001328; GB:AE001273; NID:g3329015; PIDN:AAC681
A;/Experimental source: serotype D, strain UW-3/Cx
C;/Genetics:
A;/Gene: CT578
C;/Superfamily: conserved hypothetical protein TC0867

Query Match          72.4%; Score 55; DB 2; Length 487;
Best Local Similarity 76.9%; Pred. No. 0.93;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPR 13
        |||
Db       161 ETPGLPKPGSTPR 173

RESULT 5
F81655
conserved hypothetical protein TC0867 [imported] - Chlamydia muridarum (strain Nigg)
C:/Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C:/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:/Accession: F81655
R;/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberger, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A;/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;/Reference number: A81500; MUID:20150255; PMID:10684935
A;/Accession: F81655
A;/Status: preliminary
A;/Molecule type: DNA
A;/Residues: 1-491 <TET>
A;/Cross-references: UNIPROT:Q9PJG2; GB:AE002353; GB:AE002160; NID:g7190891; PIDN:AAF396
A;/Experimental source: strain Nigg (MoPn)
C;/Genetics:
A;/Gene: TC0867
C;/Superfamily: conserved hypothetical protein TC0867

Query Match          65.8%; Score 50; DB 2; Length 491;
Best Local Similarity 69.2%; Pred. No. 5.4;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPR 13
```

Db 161 DTPGLPKPTTTPR 173

RESULT 6

S55377

urPAB protein precursor - Peptostreptococcus magnus

C;Species: Peptostreptococcus magnus

C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004

C;Accession: S55377

R;de Chateau, M.; Bjoerck, L.

submitted to the EMBL Data Library, April 1995

A;Description: Protein urPAB.

A;Reference number: S55377

A;Accession: S55377

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-345 <DEC>

A;Cross-references: UNIPROT:Q51932; EMBL:Z48975; NID:G854370; PIDN:CAA88839.1; PID:G8543

Query Match

56.6%; Score 43; DB 2; Length 345;

Best Local Similarity 62.5%; Pred. No. 44;

Matches 10; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 1 ETPELP----KPGVTP 12

Db 195 ETPETPKPKSKPDVTP 210

RESULT 7

G87106

conserved hypothetical protein gcpe [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: G87106

R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Duchoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq

A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21128732; PMID:11234002

A;Accession: G87106

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-392 <STO>

A;Cross-references: UNIPROT:Q9CBU5; GB:AL450380; NID:G13093381; PIDN:CAC30532.1; GSPDB:G

C;Genetics:

A;Gene: gcpe

C;Superfamily: gcpe protein

Query Match

56.6%; Score 43; DB 2; Length 392;

Best Local Similarity 63.6%; Pred. No. 50;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPR 13

Db 8 PEAPAPGLAPR 18

RESULT 8

T14272

cortactin-binding protein 1 - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C;Accession: T14272

R;Du, Y.; Weed, S.A.; Xiong, W.C.; Marshall, T.D.; Parsons, J.T.

Mol. Cell. Biol. 18, 5838-5851, 1998

A;Title: Identification of a novel cortactin SH3 domain-binding protein and its localization

A;Reference number: Z17952; MUID:98414600; PMID:9742101

A;Accession: T14272

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1252 <DUY>

A;Cross-references: EMBL:AF060116; NID:G3091151; PID:G3091152; PIDN:AAC62226.1

Query Match

56.6%; Score 43; DB 2; Length 1252;

Best Local Similarity 61.5%; Pred. No. 1.6e+02;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPR 13

Db 542 EQPLLPPTGAAPR 554

RESULT 9

C83598

hypothetical protein PA0371 [imported] - Pseudomonas aeruginosa (strain PA01)

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: C83598

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: C83598

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-495 <STO>

A;Cross-references: UNIPROT:Q916C3; GB:AE004475; GB:AE004091; NID:G9946221; PIDN:AAG037

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA0371

Query Match

55.3%; Score 42; DB 2; Length 495;

Best Local Similarity 77.8%; Pred. No. 39;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 11

Db 284 PETPKPGLT 292

RESULT 10

S36477

L2 protein - human papillomavirus type 15

C;Species: human papillomavirus type 15

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C;Accession: S36477

R;Delius, H.; Hofmann, B.

submitted to the EMBL Data Library, August 1993

A;Description: Primer-directed sequencing of human papillomavirus types.

A;Reference number: S36469

A;Accession: S36477

A;Molecule type: DNA

A;Residues: 1-533

A;Cross-references: UNIPROT:P36750; EMBL:X74468; NID:G396924; PIDN:CAA52510.1; PID:G396

C;Superfamily: papillomavirus L2 protein

C;Keywords: late protein

Query Match

55.3%; Score 42; DB 2; Length 533;

Best Local Similarity 63.6%; Pred. No. 96;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 TPELPKPGVTP 12

Db 83 TPTIVRPGVTP 93

RESULT 11

VGBE18

glycoprotein E - human herpesvirus 1

N;Alternate names: US8

C;Species: human herpesvirus 1

C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C;Accession: A03733; A45696
R;McGeoch, D.J.; Dolan, A.; Donald, S.; Rixon, F.J.
J. Mol. Biol. 181, 1-13, 1985
A;Title: Sequence determination and genetic content of the short unique region in the ge
A;Reference number: A00656; MUID:85160822; PMID:2984429
A;Accession: A03733
A;Molecule type: DNA
A;Residues: 1-550 <MCG>
A;Cross-references: UNIPROT:P04488; GB:X02138; NID:g59865; PIDN:CAA26062.1; PID:g59882
A;Experimental source: strain 17
R;Georgopoulou, U.; Michaelidou, A.; Roizman, B.; Mavromara-Nazos, P.
J. Virol. 67, 3961-3968, 1993
A;Title: Identification of a new transcriptional unit that yields a gene product within
A;Reference number: A45696; MUID:93287213; PMID:8389914
A;Accession: A45696
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 438-550 <GEO>
A;Cross-references: GB:S62895; NID:g386127; PIDN:AAB27080.1; PID:g386128
A;Experimental source: R35
A;Note: sequence extracted from NCBI backbone (NCBIN:133646, NCBI:P.133647)
C;Superfamily: herpesvirus glycoprotein E
C;Keywords: glycoprotein
F;124,243,501/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 55.3%; Score 42; DB 1; Length 550;
Best Local Similarity 61.5%; Pred. No. 99;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 2 TPPLPKPGVTPRS 14
||| ||| |||
Db 197 TPRLPPPAPPRS 209

RESULT 12
T34901
probable gamma-glutamyltransferase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34901
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z21561
A;Accession: T34901
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-603 <OLI>
A;Cross-references: UNIPROT:O69335; EMBL:AL023861; PIDN:CA119618.1; GSPDB:GN00070; SCOE
C;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SC3C8.26
C;Superfamily: gamma-glutamyltransferase

Query Match 55.3%; Score 42; DB 2; Length 603;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 3 PELPKPGVTPRS 14
| | | | | | | |
Db 466 PNLPGPGKRP RS 477

RESULT 13
FCSOAG
IGA Fc receptor precursor - Streptococcus agalactiae
N;Alternate names: beta antigen
C;Species: Streptococcus agalactiae
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: S15330; S20240; S17038
R;Jerlstrom, P.G.; Chhatwal, G.S.; Timmis, K.N.
Mol. Microbiol. 5, 843-849, 1991
A;Title: The IGA-binding beta antigen of the c protein complex of Group B streptococci:

A;Reference number: S15330; MUID:91312121; PMID:1857207
A;Accession: S15330
A;Molecule type: DNA
A;Residues: 1-1164 <JER1>
A;Cross-references: UNIPROT:P27951; EMBL:X59771
A;Accession: S20240
A;Molecule type: protein
A;Residues: 38-48 <JER2>
R;Jerlstrom, P.G.
submitted to the EMBL Data Library, August 1991
A;Reference number: S17038
A;Accession: S17038
A;Molecule type: DNA
A;Residues: 1-914,'E',916-1164 <JER3>
A;Cross-references: EMBL:X59771; NID:g46522; PIDN:CAA42442.1; PID:g46523
C;Superfamily: IGA Fc receptor
C;Keywords: cell wall; immunoglobulin receptor; tandem repeat; transmembrane protein
F;1-37/Domain: signal sequence #status predicted <SIG>
F;38-1164/Product: IGA Fc receptor #status experimental <MAT>
F;199-438/Domain: IGA binding #status predicted <IGA1>
F;439-826/Domain: IGA binding #status predicted <IGA2>
F;827-945/Region: proline-rich repeats
F;946-1131/Domain: cell wall-spanning #status predicted <CMS>
F;1132-1159/Domain: transmembrane #status predicted <TMM>

Query Match 55.3%; Score 42; DB 1; Length 1164;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
| | : | | | |
Db 909 EAPKIPKPKPTP 920

RESULT 14
T42710
mannose receptor, macrophage - mouse
N;Alternate names: lambda lectin; phospholipase A2 receptor
C;Species: Mus musculus (house mouse)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42710
R;Wu, K.; Yuan, J.; Laaky, L.A.
J. Biol. Chem. 271, 21323-21330, 1996
A;Title: Characterization of a novel member of the macrophage mannose receptor type C 1
A;Reference number: Z22235; MUID:96355501; PMID:8702911
A;Accession: T42710
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1479 <WUK>
A;Cross-references: UNIPROT:Q64449; EMBL:U56734; NID:g1336073; PID:g1336074; PIDN:AAC52
C;Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II r
C;Keywords: membrane protein; receptor
F;186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 55.3%; Score 42; DB 2; Length 1479;
Best Local Similarity 72.7%; Pred. No. 2.7e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 TPPLPKPGVTP 12
| | | | | | | |
Db 652 TPPLPGPDPTP 662

RESULT 15
T18262
S-layer protein - Clostridium thermocellum
C;Species: Clostridium thermocellum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18262
R;Fujino, T.; Beguin, P.; Aubert, J.P.
J. Bacteriol. 175, 1891-1899, 1993
A;Title: Organization of a Clostridium thermocellum gene cluster encoding the cellulose

A/Reference number: 218847; MUID:93209931; PMID:8458832
A/Accession: T18262
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1664 <FJ>
A/Cross-references: UNIPROT:Q06852; EMBL:X67506; NID:g296879; PID:g296881; PIDN:CAA47841

Query Match 55.3%; Score 42; DB 2; Length 1664;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ETELPKPGVTP 12
|||||
Db 1374 ETEPEPTPTTP 1385

RESULT 16
S73818
hypothetical protein H91_orf216 - Mycoplasma pneumoniae (strain ATCC 29342)
C/Species: Mycoplasma pneumoniae
A/Variety: ATCC 29342
C/Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S73818
R:Himmelreich, R.; Hilbert, H.; Plegens, H.; Pirk1, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A/Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A/Reference number: S73327; MUID:97105885; PMID:8948633
A/Accession: S73818
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-216 <HIM>
A/Cross-references: UNIPROT:P75434; EMBL:AE000048; GB:U00089; NID:g1674180; PIDN:AAB9614
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C/Genetics:
A/Genetic code: SGC3

Query Match 53.9%; Score 41; DB 2; Length 216;
Best Local Similarity 46.2%; Pred. No. 55;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETELPKPGVTPR 13
:::|
Db 64 DTPDIPKPPKPPK 76

RESULT 17
H86378
protein F21J9.16 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: H86378
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: H86378
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-519 <STO>
A/Cross-references: UNIPROT:Q9FYL3; GB:AE005172; NID:g9743337; PIDN:AAF97961.1; GSPDB:GN

A/Genetics:
A/Gene: F21J9.16
A/Map position: 1

Query Match 53.9%; Score 41; DB 2; Length 519;
Best Local Similarity 87.5%; Pred. No. 1.3e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TPELPKPG 9
|||
Db 410 TPECPKPG 417

RESULT 18
B87369
long-chain-fatty-acid-CoA ligase, probable [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: B87369
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: B87369
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-530 <STO>
A/Cross-references: UNIPROT:Q9A9L4; GB:AE005673; NID:g13422244; PIDN:AAK22950.1; GSPDB: A/Genetics:
A/Gene: CC0966
C/Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 53.9%; Score 41; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PKPGVTP 12
|||||
Db 470 PKPGVTP 476

RESULT 19
F81595
hypothetical protein CP0271 [imported] - Chlamydia pneumoniae (strain AR39)
C/Species: Chlamydia pneumoniae
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C/Accession: F81595
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: F81595
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1033 <REA>
A/Cross-references: GB:AE002187; GB:AE002161; NID:g7189192; PIDN:AAF38131.1; PID:g7189192
A/Experimental source: strain AR39, HL cells
A/Genetics:
A/Gene: CP0271

Query Match 53.9%; Score 41; DB 2; Length 1033;
Best Local Similarity 60.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
|||
Db 96 PDVPKPGTTP 105

RESULT 20
G86550
hypothetical protein CPj0483 [imported] - Chlamydia pneumoniae (strain J138)
C/Species: Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: G86550
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; I

Nucleic Acids Res. 28, 2311-2314, 2000

A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A;Reference number: A86491; MUID:20330349; PMID:10871362

A;Accession: G86550

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1043 <STO>

A;Cross-references: UNIPROT:Q9Z868; GB:BA000008; NID:g8978853; PIDN:BA98689.1; GSPDB:GN

A;Experimental source: strain J138

C;Genetics:

A;Gene: CPJ0483

Query Match

Best Local Similarity 53.9%; Score 41; DB 2; Length 1043;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12

DB 106 PDVPKPGTTP 115

RESULT 21

G72073

hypothetical protein - Chlamydia pneumoniae (strain CWL029)

C;Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: G72073

R;Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A;Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A;Reference number: A72000; MUID:99206606; PMID:10192388

A;Accession: G72073

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1043 <ARN>

A;Cross-references: UNIPROT:Q9Z868; GB:AE001633; GB:AE001363; NID:g4376762; PIDN:ADD1862

A;Experimental source: strain CWL029

C;Genetics:

A;Gene: CPN0483

Query Match

Best Local Similarity 53.9%; Score 41; DB 2; Length 1043;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12

DB 106 PDVPKPGTTP 115

RESULT 22

T06608

disease resistance protein homolog F16J13.80 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T06608

R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,

submitted to the Protein Sequence Database, April 1999

A;Reference number: Z15789

A;Accession: T06608

A;Molecule type: DNA

A;Residues: 1-1219 <BEV>

A;Cross-references: UNIPROT:Q9S266; EMBL:AL049638; GSPDB:GN00062; ATSP:F16J13.80

A;Experimental source: cultivar Columbia; BAC clone F16J13

C;Genetics:

A;Gene: ATSP:F16J13.80

A;Map position: 4

A;Introns: 158/2; 536/3; 640/3; 992/3

Query Match

Best Local Similarity 53.9%; Score 41; DB 2; Length 1219;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12

DB 124 PELTFPGVTP 133

RESULT 23

S76971

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C;Species: Synechocystis sp.

A;Variety: PCC 6803

C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 02-Jun-2003

C;Accession: S76971

R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

O, K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.

A;Reference number: S74322; MUID:97061201; PMID:8905231

A;Accession: S76971

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-152 <KAN>

A;Cross-references: EMBL:D64005; GB:AB001339; NID:g1001779; PIDN:BA10663.1; PID:g10017

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C;Genetics:

A;Start codon: GTG

C;Superfamily: cell division inhibitor minC

Query Match

Best Local Similarity 52.6%; Score 40; DB 2; Length 152;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ETPELPKP 8

DB 33 ETPSLPKP 40

RESULT 24

C48652

transfer protein spdA - Streptomyces ambofaciens plasmid pSAM2

C;Species: Streptomyces ambofaciens

C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004

C;Accession: C48652; S33428

R;Hagge, J.; Pernodet, J.L.; Sezonov, G.; Gerbaud, C.; Friedmann, A.; Guerinneau, M.

J. Bacteriol. 175, 5529-5538, 1993

A;Title: Transfer functions of the conjugative integrating element pSAM2 from Streptomy

A;Reference number: A48652; MUID:93374848; PMID:8366038

A;Accession: C48652

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-224 <HAG>

A;Cross-references: UNIPROT:Q07193; EMBL:Z19593; NID:g298051; PIDN:CAA79641.1; PID:g298

C;Genetics:

A;Genome: plasmid

Query Match

Best Local Similarity 52.6%; Score 40; DB 2; Length 224;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPV 10

DB 154 EPELPAPGL 163

RESULT 25

T46233

hypothetical protein T9C5.150 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46233

R;Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Lemcke,

submitted to the Protein Sequence Database, December 1999

A;Reference number: Z23026

A;Accession: T46233

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-261 <RIE>
A:Cross-references: UNIPROT:Q9SCK3; EMBL:AL132964
A:Experimental source: cultivar Columbia; BAC clone T9C5
C:Genetics:
A:Map position: 3
A:Introns: 101/3; 140/2; 181/3; 231/2; 252/2
A:Note: T9C5.150

Query Match 52.6%; Score 40; DB 2; Length 261;
Best Local Similarity 46.2%; Pred. No. 94;
Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 TPPLPKPGVTPRS 14
 |:|:|:|:|:|:|:
Db 81 SPEMPQAGVDPQA 93

RESULT 26

D85070
hypothetical protein AT4g05600 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85070
R/anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: D85070
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <STO>
A:Cross-references: UNIPROT:Q9M0U4; GB:NC_001268; NID:g7267321; PIDN:CAB77924.1; GSPDB:C
C:Genetics:
A:Gene: AT4g05600
A:Map position: 4

Query Match 52.6%; Score 40; DB 2; Length 275;
Best Local Similarity 53.8%; Pred. No. 99;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 TPPLPKPGVTPRS 14
 |:|:|:|:|:|:|:
Db 133 SPEVPKPPTKPSS 145

RESULT 27

A45510
probable protein kinase - maize (fragment)
C:Species: Zea mays (maize)
C:Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 09-Jul-2004
C:Accession: A45510
R:Biermann, B.; Johnson, E.M.; Feldman, L.J.
Plant Physiol. 94, 1609-1615, 1990
A:Title: Characterization and distribution of a maize cDNA encoding a peptide similar to
A:Reference number: A45510
A:Accession: A45510
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-416 <BIE>
A:Cross-references: UNIPROT:Q02494; GB:M62985; NID:g168617; PIDN:AAA33509.1; PID:g168618
C:Superfamily: probable serine/threonine-specific protein kinase ATPK64; protein kinase
F;26-367/Domain: protein kinase homology <KIN>

Query Match 52.6%; Score 40; DB 2; Length 416;
Best Local Similarity 58.3%; Pred. No. 1.5e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPRS 14
 |:|:|:|:|:|:|:
Db 381 PDIPKPEIPIRS 392

RESULT 28

T28717
hypothetical protein F10D2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T28717
R:Graves, T.; Wohlmann, P.; Gillam, B.
submitted to the EMBL Data Library, September 1997
A:Description: The sequence of C. elegans cosmid F10D2.
A:Reference number: Z20515
A:Accession: T28717
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-495 <GRA>
A:Cross-references: UNIPROT:O16912; EMBL:AF022972; PIDN:AAC48234.1; GSPDB:GN00023; CESP
A:Experimental source: strain Bristol N2; clone F10D2
C:Genetics:
A:Gene: CESP:F10D2.3
A:Map position: 5
A:Introns: 37/2; 90/1; 113/1; 183/3; 356/3; 380/3; 428/2

Query Match 52.6%; Score 40; DB 2; Length 495;
Best Local Similarity 58.3%; Pred. No. 1.8e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 PELPKPGVTPRS 14
 |:|:|:|:|:|:|:
Db 461 PSSRPPTATPRS 472

RESULT 29

A25684
hypothetical protein 108 (transposable element hobo) - fruit fly (Drosophila melanogaster
C:Species: Drosophila melanogaster
C:Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 09-Jul-2004
C:Accession: A25684
R:Streck, R.D.; MacGaffey, J.E.; Beckendorf, S.K.
EMBO J. 5, 3615-3623, 1986
A:Title: The structure of hobo transposable elements and their insertion sites.
A:Reference number: A25684
A:Accession: A25684
A:Molecule type: DNA
A:Residues: 1-644 <STR>
A:Cross-references: UNIPROT:P12258
C:Genetics:
A:Gene: FlyBase:hobo
A:Cross-references: FlyBase:FBgn0001210

Query Match 52.6%; Score 40; DB 2; Length 644;
Best Local Similarity 57.1%; Pred. No. 2.3e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ETPPLPKPGVTPRS 14
 |:|:|:|:|:|:|:
Db 538 ETPETPETPETPES 551

RESULT 30

S43768
transcription activator Vp1 - rice
C:Species: Oryza sativa (rice)
C:Date: 10-Dec-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: S43768; S48899
R:Hattori, T.; Terada, T.; Hamasuna, S.T.
Plant Mol. Biol. 24, 805-810, 1994
A:Title: Sequence and functional analyses of the rice gene homologous to the maize Vp1.
A:Reference number: S43768; MUID:94250843; PMID:8193305
A:Accession: S43768
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-728 <HAT>
A:Cross-references: UNIPROT:Q8S0A6; EMBL:D16640

A;Experimental source: strain Nipponbare
R;Hatori, T.; Terada, T.; Hamasuna, S.
submitted to the EMBL Data Library, July 1993
A;Description: Sequence and functional analysis of the rice gene homologous to the maize
A;Reference number: S48899
A;Accession: S48899
A;Molecule type: DNA
A;Residues: 1-670,'P',672-683,'P',685-728 <HAW>
A;Cross-references: EMBL:D16640; NID:g391884; PIDN:BAA04066.1; PID:g391885
A;Experimental source: strain Nipponbare
C;Genetics:
A;Gene: VP1
A;Introns: 526/3; 556/3; 590/2; 606/1; 631/3
C;Superfamily: rice transcription factor VP1
C;Keywords: DNA binding; transcription factor

Query Match 52.6%; Score 40; DB 2; Length 728;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
|:|:|:|:|
Db 263 PQLPSPGANP 272

RESULT 31
A47521
capsid protein - giardavirus GLV
C;Species: giardavirus, GLV
C;Date: 21-Jan-1994 #sequence_revision 13-Feb-1998 #text_change 16-Jul-1999
C;Accession: A47521
R;Wang, A.L.; Yang, H.M.; Shen, K.A.; Wang, C.C.
Proc. Natl. Acad. Sci. U.S.A. 90, 8595-8599, 1993
A;Title: Giardavirus double-stranded RNA genome encodes a capsid polypeptide and a gag-
A;Reference number: A47521; MUID:93391401; PMID:8378334
A;Contents: host Giardia lamblia
A;Accession: A47521
A;Molecule type: genomic RNA
A;Residues: 1-886 <WAN>
A;Cross-references: GB:L13218; NID:g1352866; PIDN:AAB01578.1; PID:g1352867
A;Note: sequence modified after extraction from NCBI backbone (NCBIN:137593, NCBI:P:13759
C;Superfamily: giardavirus capsid protein
C;Keywords: capsid protein

Query Match 52.6%; Score 40; DB 1; Length 886;
Best Local Similarity 50.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|:|:|:|:|
Db 53 ESPEVPKASIAIP 64

RESULT 32
T19140
hypothetical protein C09G5.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19140
R;Palmer, S.
submitted to the EMBL Data Library, November 1994
A;Reference number: Z19080
A;Accession: T19140
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-963 <WIL>
A;Cross-references: UNIPROT:Q09457; EMBL:Z46791; PIDN:CAA86755.1; GSPDB:GN00020; CESP:C
C;Experimental source: clone C09G5
C;Genetics:
A;Gene: CESP:C09G5.6
A;Map position: 2
A;Introns: 48/3; 862/3; 898/1

Query Match 52.6%; Score 40; DB 2; Length 963;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPR 13
|:|:|:|:|
Db 365 PQRPRTGRRR 375

RESULT 33
C47521
gag-pol-like fusion protein - giardavirus GLV
N;Alternate names: capsid protein / RNA-dependent RNA polymerase fusion protein
C;Species: giardavirus, GLV
C;Date: 06-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004
C;Accession: C47521
R;Wang, A.L.; Yang, H.M.; Shen, K.A.; Wang, C.C.
Proc. Natl. Acad. Sci. U.S.A. 90, 8595-8599, 1993
A;Title: Giardavirus double-stranded RNA genome encodes a capsid polypeptide and a gag
A;Reference number: A47521; MUID:93391401; PMID:8378334
A;Contents: host Giardia lamblia
A;Accession: C47521
A;Molecule type: genomic RNA
A;Residues: 1-1870 <FUS>
A;Cross-references: UNIPROT:Q67653; GB:L13218; NID:g1352866
C;Comment: This protein is expressed as a 190 kDa fusion protein of an N-terminal capsid
C;Keywords: translational frameshift
F;1-824/Domain: gag-like <GAG>
F;823-825/Region: minus-one translational frameshift
F;825-1870/Domain: pol-like <POL>

Query Match 52.6%; Score 40; DB 2; Length 1870;
Best Local Similarity 50.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
|:|:|:|:|
Db 53 ESPEVPKASIAIP 64

RESULT 34
T42977
large tegument protein - ateline herpesvirus 3 (strain 73)
C;Species: ateline herpesvirus 3
A;Variety: strain 73
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T42977
R;Albrecht, J.C.; Fleckenstein, B.
submitted to the EMBL Data Library, August 1998
A;Description: Primary structure of the herpesvirus ateles genome.
A;Reference number: Z22274
A;Accession: T42977
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2471 <ALB>
A;Cross-references: UNIPROT:Q9YTK3; EMBL:AF083424; PIDN:AAC95588.1
A;Experimental source: strain 73

Query Match 52.6%; Score 40; DB 2; Length 2471;
Best Local Similarity 63.6%; Pred. No. 9.1e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 TPPELPKPGVTP 12
|:|:|:|:|
Db 256 TPSPKPKSXTPT 266

RESULT 35
T03455
ALR proteain - human
C;Species: Homo sapiens (man)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004

C/Accession: T03455
R/Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, Oncogene 15, 549-560, 1997
A/Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
A/Reference number: Z14954; MUID:97388474; PMID:9247308
A/Accession: T03455
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-4957 <PRA>
A/Cross-references: UNIPROT:O14686; EMBL:AF010404; NID:g2358286; PIDN:AAC51735.1; PID:g2
C/Genetics:
A/Gene: ALR
A/Map position: 12
C/Superfamily: acute lymphoblastic leukemia protein, ALR type
C/Keywords: alternative splicing

Query Match 52.6%; Score 40; DB 2; Length 4957;
Best Local Similarity 58.3%; Pred. No. 1.8e+03;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 TPPLPKPGVTPR 13
:|||||:|:
Db 272 SPELEKPLSPR 283

RESULT 36
T12738
ATP/GTP-binding protein 22 - Methanobacterium phage psiM2
C/Species: Methanobacterium phage psiM2
C/Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C/Accession: T12738
R/Pfister, P.; Wasserfallen, A.; Stettler, R.; Leisinger, T.
Submitted to the EMBL Data Library, May 1998
A/Description: Archaeophage PsiM2 complete genomic DNA.
A/Reference number: Z17578
A/Accession: T12738
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-144 <PFI>
A/Cross-references: UNIPROT:O80212; EMBL:AF065411; NID:g3249585; PID:g3249607; PIDN:AAC2
A/Experimental source: host Methanobacterium thermoautotrophicum strain Marburg
C/Superfamily: Methanobacterium phage psiM2 hypothetical ATP/GTP-binding protein 22
C/Keywords: nucleotide binding; P-loop
F/34-41/Region: nucleotide-binding motif A (P-loop)

Query Match 51.3%; Score 39; DB 2; Length 144;
Best Local Similarity 87.5%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 LPKPGVTP 12
|||||:
Db 35 LPKPGKTP 42

RESULT 37
T17931
IGA Fc receptor-like protein A428L - Chlorella virus PBCV-1
C/Species: Chlorella virus PBCV-1
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T17931
R/Graves, M.V.; Van Etten, J.L.
Submitted to the EMBL Data Library, May 1999
A/Reference number: Z18806
A/Accession: T17931
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-145 <GRA>
A/Cross-references: UNIPROT:Q98480; EMBL:U42580; NID:g4028896; PIDN:AAC96796.1
A/Experimental source: specific host Chlorella strain NC64A
C/Genetics:
A/Note: A428L

Query Match 51.3%; Score 39; DB 2; Length 145;

Best Local Similarity 58.3%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ETPPLPKPGVTP 12
|||||:
Db 77 ETPPIPSVPTTP 88

RESULT 38
B26414
95K nonspecific cross-reacting antigen - human (fragments)
C/Species: Homo sapiens (man)
C/Date: 11-May-1989 #sequence_revision 20-Oct-1989 #text_change 16-Jul-1999
C/Accession: B26414
R/Paxton, R.J.; Mooser, G.; Pande, H.; Lee, T.D.; Shively, J.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 920-924, 1987
A/Title: Sequence analysis of carcinoembryonic antigen: identification of glycosylation
A/Reference number: A26414; MUID:87147209; PMID:3469650
A/Accession: B26414
A/Status: preliminary
A/Molecule type: protein
A/Residues: 1-172 <PAX>
C/Superfamily: carcinoembryonic antigen; carcinoembryonic antigen precursor amino-termi
F/135-165/Domain: immunoglobulin homology <IMM>

Query Match 51.3%; Score 39; DB 2; Length 172;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVT 11
|||||:
Db 80 PELPKPSIS 88

RESULT 39
C40428
nonspecific cross-reacting antigen W282 precursor - human
N/Alternate names: carcinoembryonic antigen homolog CGM1 (version 3)
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C/Accession: C40428; D44476
R/Kuroki, M.; Arakawa, F.; Matsuo, Y.; Oikawa, S.; Misumi, Y.; Nakazato, H.; Matsuo, J.
J. Biol. Chem. 266, 11810-11817, 1991
A/Title: Molecular cloning of nonspecific cross-reacting antigens in human granulocytes
A/Reference number: A40428; MUID:91268052; PMID:2050678
A/Accession: C40428
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-177 <KUR>
A/Cross-references: GB:D90278; NID:g219536; PIDN:BAI4322.1; PID:g219537
R/Khan, W.N.; Fraengsmyr, L.; Teglund, S.; Israelsson, A.; Bremer, K.; Hammarstroem, S.
Genomics 14, 384-390, 1992
A/Title: Identification of three new genes and estimation of the size of the carcinoemb
A/Reference number: A44476; MUID:93052339; PMID:1427854
A/Accession: D44476
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 35-141 <KHA>
C/Genetics:
A/Gene: GDB:NGA
C/Superfamily: nonspecific cross-reacting antigen; carcinoembryonic antigen precursor a
F/1-138/Domain: carcinoembryonic antigen precursor amino-terminal homology <CEAN>

Query Match 51.3%; Score 39; DB 1; Length 177;
Best Local Similarity 66.7%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVT 11
|||||:
Db 142 PELPKPSIS 150

RESULT 40

F83453
adenine phosphoribosyltransferase PA1543 [imported] - Pseudomonas aeruginosa (strain PAO
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83453
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
. ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83453
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-182 <STO>
A;Cross-references: UNIPROT:Q04633; GB:AE004582; GB:AE004091; NID:g9947492; PIDD:AA0493
A;Experimental source: strain PA01
C;Genetics:
A;Gene: apt; PA1543
C;Superfamily: adenine phosphoribosyltransferase

Query Match 51.3%; Score 39; DB 2; Length 182;
Best local similarity 63.6%; Pred. No. 92;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPR 13
|:|||||
Db 16 PDFPKPGVFR 26

Search completed: October 28, 2005, 22:18:30
Job time : 32.6087 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 22:15:59 ; Search time 139.391 Seconds
(without alignments)
41.987 Million cell updates/sec

Title: US-10-634-914-16

Perfect score: 76

Sequence: 1 ETPELPKPGVTPRS 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1865214 seqs, 418043040 residues

Total number of hits satisfying chosen parameters: 1865214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*

19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep:*

20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*

21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	14	US-10-634-914-16	Sequence 16, Appl
2	76	100.0	302	US-10-634-914-3	Sequence 3, Appl
3	76	100.0	493	US-09-820-843A-51	Sequence 51, Appl
4	76	100.0	496	US-10-289-762-867	Sequence 867, App
5	76	100.0	496	US-10-634-914-2	Sequence 2, Appl
6	72.4	487	15	US-10-282-122A-55018	Sequence 55018, A
7	50	65.8	535	US-10-450-763-33751	Sequence 33751, A
8	47	61.8	181	US-10-156-761-12554	Sequence 12554, A
9	47	61.8	228	US-10-450-763-45435	Sequence 45435, A
10	47	61.8	267	US-10-739-930-9579	Sequence 9579, App
11	47	61.8	1566	US-10-491-545A-63	Sequence 63, Appl

12	46	60.5	133	US-10-450-763-50962	Sequence 50962, A
13	46	60.5	133	US-10-450-763-54305	Sequence 54305, A
14	46	60.5	181	US-10-450-763-40067	Sequence 40067, A
15	46	60.5	189	US-10-425-115-227818	Sequence 227818, A
16	46	60.5	258	US-10-425-114-67973	Sequence 67973, A
17	46	60.5	258	US-10-425-115-227817	Sequence 227817, A
18	46	60.5	259	US-10-425-115-207691	Sequence 207691, A
19	46	60.5	278	US-10-425-115-227821	Sequence 227821, A
20	46	60.5	283	US-10-425-114-67521	Sequence 67521, A
21	46	60.5	292	US-10-450-763-37181	Sequence 37181, A
22	46	60.5	294	US-10-450-763-50757	Sequence 50757, A
23	46	60.5	309	US-10-450-763-40062	Sequence 40062, A
24	46	60.5	309	US-10-450-763-50949	Sequence 50949, A
25	46	60.5	310	US-10-450-763-31184	Sequence 31184, A
26	46	60.5	310	US-10-450-763-33421	Sequence 33421, A
27	46	60.5	310	US-10-450-763-40060	Sequence 40060, A
28	46	60.5	310	US-10-450-763-48965	Sequence 48965, A
29	46	60.5	310	US-10-450-763-50948	Sequence 50948, A
30	46	60.5	310	US-10-450-763-56923	Sequence 56923, A
31	46	60.5	485	US-09-808-483-12	Sequence 12, Appl
32	46	60.5	485	US-11-097-143-16572	Sequence 16572, A
33	46	60.5	531	US-10-425-114-70256	Sequence 70256, A
34	46	60.5	535	US-09-808-483-10	Sequence 10, Appl
35	45	59.2	95	US-10-450-763-40028	Sequence 40028, A
36	45	59.2	156	US-10-767-701-32834	Sequence 32834, A
37	45	59.2	158	US-10-450-763-39062	Sequence 39062, A
38	45	59.2	172	US-10-767-701-59998	Sequence 59998, A
39	45	59.2	251	US-10-425-115-207693	Sequence 207693, A
40	45	59.2	316	US-10-437-963-111175	Sequence 111175, A
41	44	57.9	184	US-10-767-701-42472	Sequence 42472, A
42	44	57.9	200	US-10-450-763-37900	Sequence 37900, A
43	44	57.9	241	US-10-259-165-176	Sequence 176, App
44	44	57.9	241	US-10-437-963-176969	Sequence 176969, A
45	44	57.9	241	US-10-732-923-5688	Sequence 5688, App
46	44	57.9	299	US-10-094-749-1848	Sequence 1848, App
47	44	57.9	313	US-10-282-122A-55940	Sequence 55940, A
48	44	57.9	1241	US-10-494-836-62	Sequence 62, Appl
49	44	57.9	1257	US-09-738-626-4750	Sequence 4750, App
50	44	57.9	2426	US-10-764-425-172	Sequence 172, App
51	44	57.9	2447	US-10-450-763-52739	Sequence 52739, A
52	43.5	57.2	415	US-10-739-930-9586	Sequence 9586, App
53	43.5	57.2	419	US-09-934-455-422	Sequence 422, App
54	43.5	57.2	419	US-10-412-639B-738	Sequence 738, App
55	43	56.6	58	US-09-864-761-36383	Sequence 36383, A
56	43	56.6	59	US-10-437-963-162875	Sequence 162875, A
57	43	56.6	60	US-10-425-115-244529	Sequence 244529, A
58	43	56.6	69	US-10-767-701-58805	Sequence 58805, A
59	43	56.6	82	US-10-425-115-340385	Sequence 340385, A
60	43	56.6	119	US-10-425-115-330323	Sequence 330323, A
61	43	56.6	172	US-10-424-599-281384	Sequence 281384, A
62	43	56.6	175	US-10-425-115-217317	Sequence 217317, A
63	43	56.6	189	US-10-425-114-69001	Sequence 69001, A
64	43	56.6	235	US-10-424-599-268760	Sequence 268760, A
65	43	56.6	267	US-10-425-115-208027	Sequence 208027, A
66	43	56.6	309	US-10-425-115-208032	Sequence 208032, A
67	43	56.6	392	US-10-282-122A-63883	Sequence 63883, A
68	43	56.6	435	US-10-437-963-112015	Sequence 112015, A
69	43	56.6	2951	US-11-097-143-7665	Sequence 7665, App
70	42	55.3	67	US-09-814-122-45	Sequence 45, Appl
71	42	55.3	71	US-10-649-857-45	Sequence 45, Appl
72	42	55.3	78	US-09-801-968-29	Sequence 29, Appl
73	42	55.3	78	US-09-801-968-30	Sequence 30, Appl
74	42	55.3	78	US-09-802-154-29	Sequence 29, Appl
75	42	55.3	78	US-09-802-154-30	Sequence 30, Appl
76	42	55.3	115	US-10-425-115-207696	Sequence 207696, A
77	42	55.3	144	US-10-437-963-174088	Sequence 174088, A
78	42	55.3	198	US-09-925-300-1562	Sequence 1562, App
79	42	55.3	198	US-10-437-963-123573	Sequence 123573, A
80	42	55.3	247	US-09-822-485-17	Sequence 17, Appl
81	42	55.3	247	US-09-750-963-12	Sequence 12, Appl
82	42	55.3	247	US-09-989-722-499	Sequence 499, App
83	42	55.3	247	US-09-989-723-499	Sequence 499, App
84	42	55.3	247	US-09-989-279-499	Sequence 499, App

```
85 42 55.3 247 9 US-09-989-727-499 Sequence 499, App
86 42 55.3 247 9 US-09-989-731-499 Sequence 499, App
87 42 55.3 247 9 US-09-989-732-499 Sequence 499, App
88 42 55.3 247 9 US-09-991-073-499 Sequence 499, App
89 42 55.3 247 9 US-09-990-442-499 Sequence 499, App
90 42 55.3 247 9 US-09-991-163-499 Sequence 499, App
91 42 55.3 247 9 US-09-993-604-499 Sequence 499, App
92 42 55.3 247 9 US-09-990-456-499 Sequence 499, App
93 42 55.3 247 9 US-09-989-721-499 Sequence 499, App
94 42 55.3 247 9 US-09-989-721-499 Sequence 499, App
95 42 55.3 247 9 US-09-989-293A-499 Sequence 499, App
96 42 55.3 247 9 US-09-989-735-499 Sequence 499, App
97 42 55.3 247 9 US-09-990-444-499 Sequence 499, App
98 42 55.3 247 9 US-09-991-181-499 Sequence 499, App
99 42 55.3 247 9 US-09-989-730-499 Sequence 499, App
100 42 55.3 247 9 US-09-990-436-499 Sequence 499, App
```

ALIGNMENTS

RESULT 1

```
US-10-634-914-16
; Sequence 16, Application US/10634914
; Publication No. US20040029806A1
; GENERAL INFORMATION:
; APPLICANT: BURNIE, JAMES PETER
; APPLICANT: MATTHEWS, RUTH CHRISTINE
; TITLE OF INVENTION: MEDICAMENT
; FILE REFERENCE: 050885-0281578
; CURRENT APPLICATION NUMBER: US/10/634,914
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: US/09/889,314
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: GB 9902555.3
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: PCT/GB00/00237
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-634-914-16
```

```
Query Match 100.0%; Score 76; DB 15; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 1 ETPELPKPGVTPRS 14
```

RESULT 2

```
US-10-634-914-3
; Sequence 3, Application US/10634914
; Publication No. US20040029806A1
; GENERAL INFORMATION:
; APPLICANT: BURNIE, JAMES PETER
; APPLICANT: MATTHEWS, RUTH CHRISTINE
; TITLE OF INVENTION: MEDICAMENT
; FILE REFERENCE: 050885-0281578
; CURRENT APPLICATION NUMBER: US/10/634,914
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: US/09/889,314
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: GB 9902555.3
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: PCT/GB00/00237
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 16
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Codon
; OTHER INFORMATION: optimised N-terminal section of Chlamydia
; OTHER INFORMATION: pneumoniae protein
; NAME/KEY: UNSURE
; LOCATION: (1)..(30)
; OTHER INFORMATION: S-tag and thrombin cleavage site
; FEATURE:
; OTHER INFORMATION: Positions (297)..(302) comprise Histidine tag
US-10-634-914-3
```

```
Query Match 100.0%; Score 76; DB 15; Length 302;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 199 ETPELPKPGVTPRS 212
```

RESULT 3

```
US-09-820-843A-51
; Sequence 51, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTE
; FILE REFERENCE: Q63915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
; LENGTH: 493
; TYPE: PRT
; ORGANISM: C. pneumoniae CWL029
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: CT578 hypothetical protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: gi|4377121
US-09-820-843A-51
```

```
Query Match 100.0%; Score 76; DB 10; Length 493;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 164 ETPELPKPGVTPRS 177
```

RESULT 4

```
US-10-289-762-867
; Sequence 867, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Grifflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragment
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prev
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 867
; LENGTH: 496
```



```
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-289-762-867

Query Match          100.0%; Score 76; DB 15; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||
Db      167 ETPELPKPGVTPRS 180

RESULT 5
US-10-634-914-2
; Sequence 2, Application US/10634914
; Publication No. US20040029806A1
; GENERAL INFORMATION:
; APPLICANT: BURNIE, JAMES PETER
; APPLICANT: MATTHEWS, RUTH CHRISTINE
; TITLE OF INVENTION: MEDICAMENT
; FILE REFERENCE: 050885-0281578
; CURRENT APPLICATION NUMBER: US/10/634,914
; CURRENT FILING DATE: 2003-08-06
; PRIOR APPLICATION NUMBER: US/09/889,314
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: GB 9902555.3
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: PCT/GB00/00237
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-10-634-914-2

Query Match          100.0%; Score 76; DB 15; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        |||||
Db      167 ETPELPKPGVTPRS 180

RESULT 6
US-10-282-122A-55018
; Sequence 55018, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
```

```
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55018
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Chlamydia trachomatis
US-10-282-122A-55018

Query Match          72.4%; Score 55; DB 15; Length 487;
Best Local Similarity 76.9%; Pred. No. .25;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPR 13
        |||||
Db      161 ETPGLPKPSTTPR 173

RESULT 7
US-10-450-763-33751
; Sequence 33751, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 33751
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (366)..(379)
; OTHER INFORMATION: LUTBOVIRUS ORF6 PROTEIN SIGNATURE domain identified by
; OTHER INFORMATION: eMATRIX, accession number PR00910A, p-value=1.000e-09, raw score
; OTHER INFORMATION: 2.51
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(535)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-33751
```

```
Query Match          65.8%; Score 50; DB 18; Length 535;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTPRS 14
        :||:|||||:|
```

Db 178 QTPSVPKPGLEPTS 191

RESULT 8
US-10-156-761-12554
; Sequence 12554, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12554
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12554

Query Match 61.8%; Score 47; DB 14; Length 181;
Best Local Similarity 61.5%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ETPELPKPGVTPR 13
Db 167 EEPERPRPGSSPR 179

RESULT 9
US-10-450-763-45435
; Sequence 45435, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 45435
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(228)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-45435

Query Match 61.8%; Score 47; DB 18; Length 228;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ETPELPKPGVTP 12
:|||||:|||||

Db 28 QTPELQRPQVPP 39

RESULT 10
US-10-739-930-9579
; Sequence 9579, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9579
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-C119372_1.p
US-10-739-930-9579

Query Match 61.8%; Score 47; DB 16; Length 267;
Best Local Similarity 57.1%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ETPELPKPGVTPRS 14
Db 36 ETPAAPAGGTPPKA 49

RESULT 11
US-10-491-545A-63
; Sequence 63, Application US/10491545A
; Publication No. US20050130117A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Masuda, Esteban
; APPLICANT: Pardo, Jorge
; APPLICANT: Zhao, Haoran
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Lymphocyte Activation and Migration
; FILE REFERENCE: 021044-000330US
; CURRENT APPLICATION NUMBER: US/10/491,545A
; CURRENT FILING DATE: 2004-04-01
; PRIOR APPLICATION NUMBER: US 60/327,212
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: WO PCT/US02/31618
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 1566
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human RERE
US-10-491-545A-63

Query Match 61.8%; Score 47; DB 18; Length 1566;
Best Local Similarity 72.7%; Pred. No. 9.6e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TPELPKPGVTP 12
Db 758 TPQLPTPGTTP 768

RESULT 12
US-10-450-763-50962

```
; Sequence 50962, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50962
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-50962
```

```
Query Match      60.5%; Score 46; DB 18; Length 133;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTP 12
      :|||||:|||||
Db      52 QTEPQQPGVPP 63
```

```
RESULT 13
US-10-450-763-54305
; Sequence 54305, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 54305
; LENGTH: 133
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-54305
```

```
Query Match      60.5%; Score 46; DB 18; Length 133;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTP 12
      :|||||:|||||
Db      52 QTEPQQPGVPP 63
```

```
RESULT 14
US-10-450-763-40067
; Sequence 40067, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
```

```
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40067
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-40067
```

```
Query Match      60.5%; Score 46; DB 18; Length 181;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1 ETPELPKPGVTP 12
      :|||||:|||||
Db      100 QTEPQQPGVPP 111
```

```
RESULT 15
US-10-425-115-227818
; Sequence 227818, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 227818
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(189)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_139366C.1.pep
US-10-425-115-227818
```

```
Query Match      60.5%; Score 46; DB 16; Length 189;
Best Local Similarity 80.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      3 PELPKPGVTP 12
      :|||||:|||||
Db      132 PELPKPELTP 141
```

```
RESULT 16
US-10-425-114-67973
; Sequence 67973, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67973
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE047H09_FLI.pep
US-10-425-114-67973

Query Match 60.5%; Score 46; DB 15; Length 258;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPRS 14
DB 88 PELPKPEIPPHS 99

RESULT 17
US-10-425-115-227817
; Sequence 227817, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 227817
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_139365C.1.pep
US-10-425-115-227817

Query Match 60.5%; Score 46; DB 16; Length 258;
Best Local Similarity 80.0%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
DB 159 PELPKPELTP 168

RESULT 18
US-10-425-115-207691
; Sequence 207691, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 207691
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:

; OTHER INFORMATION: Clone ID: MRT4577_121003C.1.pep
US-10-425-115-207691
Query Match 60.5%; Score 46; DB 16; Length 259;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPRS 14
DB 89 PELPKPEIPPHS 100

RESULT 19
US-10-425-115-227821
; Sequence 227821, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 227821
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_139369C.1.pep
US-10-425-115-227821

Query Match 60.5%; Score 46; DB 16; Length 278;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
DB 145 PELPKPELTP 154

RESULT 20
US-10-425-114-67521
; Sequence 67521, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 67521
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-005-B9_FLI.pep
US-10-425-114-67521

Query Match 60.5%; Score 46; DB 15; Length 283;
Best Local Similarity 80.0%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVTP 12
:|||||:|
Db 150 PELPKPELTP 159

RESULT 21

US-10-450-763-37181
; Sequence 37181, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 37181
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-37181

Query Match 60.5%; Score 46; DB 18; Length 292;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
:|||||:|
Db 211 QTPELQPGVPP 222

RESULT 22

US-10-450-763-50757
; Sequence 50757, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50757
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-50757

Query Match 60.5%; Score 46; DB 18; Length 294;
Best Local Similarity 66.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
:|||||:|
Db 213 QTPELQPGVPP 224

RESULT 23
US-10-450-763-40062

; Sequence 40062, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40062
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-40062

Query Match 60.5%; Score 46; DB 18; Length 309;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
:|||||:|
Db 228 QTPELQPGVPP 239

RESULT 24

US-10-450-763-50949
; Sequence 50949, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50949
; LENGTH: 309
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-50949

Query Match 60.5%; Score 46; DB 18; Length 309;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
:|||||:|
Db 228 QTPELQPGVPP 239

RESULT 25

US-10-450-763-31184
; Sequence 31184, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763

```
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 31184
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-31184

Query Match          60.5%; Score 46; DB 18; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTP 12
       :|||||:|||||
Db      229 QTPELQQPGVPP 240

RESULT 26
US-10-450-763-33421
; Sequence 33421, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 33421
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-33421

Query Match          60.5%; Score 46; DB 18; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTP 12
       :|||||:|||||
Db      229 QTPELQQPGVPP 240

RESULT 27
US-10-450-763-40060
; Sequence 40060, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
```

```
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40060
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-40060

Query Match          60.5%; Score 46; DB 18; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTP 12
       :|||||:|||||
Db      229 QTPELQQPGVPP 240

RESULT 28
US-10-450-763-48965
; Sequence 48965, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 48965
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-48965

Query Match          60.5%; Score 46; DB 18; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 ETPELPKPGVTP 12
       :|||||:|||||
Db      229 QTPELQQPGVPP 240

RESULT 29
US-10-450-763-50948
; Sequence 50948, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 50948
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-50948
```

Query Match 60.5%; Score 46; DB 18; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETELPKPGVTP 12
:||||:||||
Db 229 QTEPQQPGVPP 240

RESULT 30

US-10-450-763-56923
; Sequence 56923, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; PRIOR FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 56923
; LENGTH: 310
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-450-763-56923

Query Match 60.5%; Score 46; DB 18; Length 310;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETELPKPGVTP 12
:||||:||||
Db 229 QTEPQQPGVPP 240

RESULT 31

US-09-808-483-12
; Sequence 12, Application US/09808483
; Patent No. US20020001824A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Ligand-gated anion channels of insects
; FILE REFERENCE: Le A 34 397
; CURRENT APPLICATION NUMBER: US/09/808,483
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: DE 100 136 19.2
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 12
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-808-483-12

Query Match 60.5%; Score 46; DB 9; Length 485;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
:||||:||||
Db 438 PELPKPKLTP 447

RESULT 32

US-11-097-143-16572
; Sequence 16572, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16572
; LENGTH: 485
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-16572

Query Match 60.5%; Score 46; DB 20; Length 485;
Best Local Similarity 80.0%; Pred. No. 4.1e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
:||||:||||
Db 438 PELPKPKLTP 447

RESULT 33

US-10-425-114-70256
; Sequence 70256, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70256
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73004G01_FLI.pep
US-10-425-114-70256

Query Match 60.5%; Score 46; DB 15; Length 531;
Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
| | | | | : |
Db 25 ETPTLPKPQIQP 36

RESULT 34

US-09-808-483-10
; Sequence 10, Application US/09808483
; Patent No. US20020001824A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Ligand-gated anion channels of insects
; FILE REFERENCE: Le A 34 397
; CURRENT APPLICATION NUMBER: US/09/808,483
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: DE 100 136 19.2
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-808-483-10

Query Match 60.5%; Score 46; DB 9; Length 535;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVTP 12
| | | | | : | |
Db 438 PELPKPKLTP 447

RESULT 35

US-10-450-763-40028
; Sequence 40028, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 40028
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (25)..(42)
; OTHER INFORMATION: Zinc finger, CCHC class domain identified by Pfam, accession
; OTHER INFORMATION: name zf-CCHC, E-value=5e-05, Pfam score of 30.1
US-10-450-763-40028

Query Match 59.2%; Score 45; DB 18; Length 95;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
: | | | | : | | | |
Db 8 KTEPLQOPGVVP 19

RESULT 36
US-10-767-701-32834
; Sequence 32834, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 32834
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(156)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C1899_1.pep
US-10-767-701-32834

Query Match 59.2%; Score 45; DB 16; Length 156;
Best Local Similarity 61.5%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPR 13
| | | | | : | |
Db 34 ELPLPKPEIIPR 46

RESULT 37

US-10-450-763-39062
; Sequence 39062, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 39062
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (89)..(106)
; OTHER INFORMATION: Zinc finger, CCHC class domain identified by Pfam, accession
; OTHER INFORMATION: name zf-CCHC, E-value=5e-05, Pfam score of 30.1
; NAME/KEY: misc_feature
; LOCATION: (1)..(158)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-39062

Query Match 59.2%; Score 45; DB 18; Length 158;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12

Db 71 KTEPQQPGVPP 82

RESULT 38

US-10-767-701-59998
; Sequence 59998, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 59998
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 7551639.pap
US-10-767-701-59998

Query Match 59.2%; Score 45; DB 16; Length 172;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETEPLPKPGVTP 12
Db 45 EVPELPKPELPP 56

RESULT 39

US-10-425-115-207693
; Sequence 207693, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 207693
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_121005C.1.pap
US-10-425-115-207693

Query Match 59.2%; Score 45; DB 16; Length 251;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETEPLPKPGVTP 12
Db 44 EVPELPKPELPP 55

RESULT 40

US-10-437-963-111175
; Sequence 111175, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 111175
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_15178C.1.pap
US-10-437-963-111175

Query Match 59.2%; Score 45; DB 16; Length 316;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 PELPKPGVTPRS 14
Db 299 PELPLPSLVPRS 310

Search completed: October 28, 2005, 22:37:30
Job time: 142.391 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2005, 21:53:40 : Search time 154 Seconds
(without alignments)
35.160 Million cell updates/sec

Title: US-10-634-914-16
Perfect score: 76
Sequence: 1 ETEPLPKGVTPRS 14

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp19808:*

2: geneseqp19908:*

3: geneseqp20008:*

4: geneseqp20018:*

5: geneseqp20028:*

6: geneseqp20038:*

7: geneseqp20038:*

8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	76	100.0	14	3	AAB08336	Aab08336 Epitope d
2	76	100.0	259	2	AAR94586	Aar94586 C. pneumo
3	76	100.0	259	2	AAW01743	Aaw01743 C. pneumo
4	76	100.0	271	2	AAR94580	Aar94580 C. pneumo
5	76	100.0	302	3	AAB08323	Aab08323 Amino aci
6	76	100.0	432	2	AAR94585	Aar94585 DHFR/C. p
7	76	100.0	488	2	AAR94579	Aar94579 Chlamydia
8	76	100.0	493	7	ABO23550	Ab023550 Chlamydia
9	76	100.0	496	2	AAV35449	Aay35449 Amino aci
10	76	100.0	496	3	AAB08322	Aab08322 Amino aci
11	76	100.0	649	2	AAR94584	Aar94584 DHFR/C. p
12	55	72.4	481	2	AAV37632	Aay37632 Protein w
13	55	72.4	487	6	ABU27094	Abu27094 Protein e
14	50	65.8	535	4	ABG03392	Abg03392 Novel hum
15	47	61.8	228	4	ABG15076	Abg15076 Novel hum
16	47	61.8	1012	2	AAV17406	Aay17406 Human atr
17	47	61.8	1566	6	ABR59716	Abri59716 Human RER
18	46	60.5	133	4	ABG20603	Abg20603 Novel hum
19	46	60.5	133	4	ABG23946	Abg23946 Novel hum
20	46	60.5	181	4	ABG09708	Abg09708 Novel hum
21	46	60.5	292	4	ABG06822	Abg06822 Novel hum
22	46	60.5	294	4	ABG20398	Abg20398 Novel hum
23	46	60.5	309	4	ABG20590	Abg20590 Novel hum
24	46	60.5	309	4	ABG09703	Abg09703 Novel hum
25	46	60.5	310	4	ABG26564	Abg26564 Novel hum

26	46	60.5	310	4	ABG09701	Abg09701 Novel hum
27	46	60.5	310	4	ABG00825	Abg00825 Novel hum
28	46	60.5	310	4	ABG20589	Abg20589 Novel hum
29	46	60.5	310	4	ABG18606	Abg18606 Novel hum
30	46	60.5	310	4	ABG03062	Abg03062 Novel hum
31	46	60.5	485	4	ABB63260	Abb63260 Drosophila
32	46	60.5	485	4	AAAG80006	Aag80006 D. melano
33	46	60.5	485	4	AAAG67566	Aag67566 Amino aci
34	46	60.5	535	4	AAAG80005	Aag80005 D. melano
35	46	60.5	2963	2	AAW56444	Aaw56444 Fragment
36	45	59.2	95	4	ABG09669	Abg09669 Novel hum
37	45	59.2	158	4	ABG08703	Abg08703 Novel hum
38	44	57.9	200	4	ABG07541	Abg07541 Novel hum
39	44	57.9	241	7	ADJ11540	Adj11540 Rice prot
40	44	57.9	299	6	ADA54280	Ada54280 Human pro
41	44	57.9	313	6	ABU28016	Abu28016 Protein e
42	44	57.9	671	7	ABO68176	Ab068176 Pseudomon
43	44	57.9	1241	7	ADD13377	Add13377 C. glutam
44	44	57.9	1257	2	AAAR87628	Aar87628 Alpha-ket
45	44	57.9	1257	2	AAW41781	Aaw41781 B. lactof
46	44	57.9	1257	4	AAAG90996	Aag90996 C. glutami
47	44	57.9	1446	8	ADR09797	Adr09797 Human pro
48	44	57.9	2447	4	ABG22380	Abg22380 Novel hum
49	43.5	57.2	419	5	AAU93147	Aau93147 Arabidops
50	43.5	57.2	419	8	ADO02325	Ado02325 Thalecres
51	43.5	57.2	419	8	ADO61715	Ado61715 Transcrip
52	43	56.6	58	4	AAAM1672	Aam1672 Peptide #
53	43	56.6	58	4	ABAB35656	Abb35656 Peptide #
54	43	56.6	58	4	AAAM29156	Aam29156 Peptide #
55	43	56.6	58	4	ABBB30489	Abb30489 Peptide #
56	43	56.6	58	4	ABBB21085	Abb21085 Protein #
57	43	56.6	58	4	AAAM68849	Aam68849 Human bon
58	43	56.6	58	4	AAAM56472	Aam56472 Human bra
59	43	56.6	58	4	ABG50508	Abg50508 Human liv
60	43	56.6	58	4	AAAM04388	Aam04388 Peptide #
61	43	56.6	58	5	ABG38430	Abg38430 Human pep
62	43	56.6	236	7	ABO78128	Ab078128 Pseudomon
63	43	56.6	261	7	ABO78753	Ab078753 Pseudomon
64	43	56.6	392	6	ABU35959	Abu35959 Protein e
65	43	56.6	731	7	ABO68434	Ab068434 Pseudomon
66	43	56.6	1470	4	ABAB31518	Abb31518 Amino aci
67	43	56.6	2951	4	ABBB60291	Abb60291 Drosophil
68	42	55.3	51	2	AAV48336	Aay48336 Human pro
69	42	55.3	67	2	AAW83944	Aaw83944 Human sec
70	42	55.3	80	4	AAAG64088	Aag64088 Fumarase
71	42	55.3	150	2	AAV60549	Aay60549 Human nor
72	42	55.3	159	7	ABO81043	Ab081043 Pseudomon
73	42	55.3	198	3	AAAB56984	Aab56984 Human pro
74	42	55.3	247	2	AAW53013	Aaw53013 Fibroblas
75	42	55.3	247	4	AAAB31182	Aab31182 Amino aci
76	42	55.3	247	4	AAU12313	Aau12313 Human PRO
77	42	55.3	247	4	AAAB47288	Aab47288 PRO185 po
78	42	55.3	247	4	AAAG65660	Aag65660 Human fib
79	42	55.3	247	4	AAAE04406	Aae04406 Human fib
80	42	55.3	247	4	AAAB65291	Aab65291 Human PRO
81	42	55.3	247	4	AAAB85825	Aab85825 Human RGF
82	42	55.3	247	5	AAAE18819	Aae18819 Human PRO
83	42	55.3	247	6	ABU58106	Abu58106 Human PRO
84	42	55.3	247	6	ABU59184	Abu59184 Novel hum
85	42	55.3	247	6	ABU82696	Abu82696 Human sec
86	42	55.3	247	6	ABO17757	Ab017757 Novel hum
87	42	55.3	247	6	ABU60615	Abu60615 Human sec
88	42	55.3	247	6	ABO25153	Ab025153 Novel hum
89	42	55.3	247	6	ABU13997	Abu13997 Human PRO
90	42	55.3	247	6	ABU81011	Abu81011 Human PRO
91	42	55.3	247	6	ABU72582	Abu72582 Novel hum
92	42	55.3	247	6	ABU66711	Abu66711 Human PRO
93	42	55.3	247	6	ABU67271	Abu67271 Novel hum
94	42	55.3	247	6	ABU59792	Abu59792 Novel sec
95	42	55.3	247	6	ABU59331	Abu59331 Human sec
96	42	55.3	247	6	ABO26028	Ab026028 Human PRO
97	42	55.3	247	6	ABO24982	Ab024982 Human sec
98	42	55.3	247	6	ABU72039	Abu72039 Novel hum

99 42 55.3 247 6 ABU67140 hum
100 42 55.3 247 6 ABU59037 sec

ALIGNMENTS

RESULT 1
AAB08336
ID AAB08336 standard; peptide; 14 AA.
XX
AC AAB08336;
XX
DT 12-SEP-2003 (revised)
DT 04-DEC-2000 (first entry)
XX
DE Epitope derived from a hlamydia pneumoniae protein.
XX
KM Chlamydia pneumoniae protein; infection; arteriosclerosis;
KM coronary atherosclerosis; epitope.
XX
OS Chlamydophila pneumoniae.
XX
PN WO200046359-A2.
XX
PD 10-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-GB000237.
XX
PR 05-FEB-1999; 99GB-00002555.
XX
PA (NEUT-) NEUTEC PHARMA PLC.
XX
PI Burnie JP, Matchews RC;
XX
DR WPI; 2000-543485/49.
XX
PT New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
PT preventing and treating C. pneumoniae infection and arteriosclerosis,
PT including coronary atherosclerosis.
XX
PS Claim 5; Page 35; 35pp; English.
XX
CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
CC protein. The protein, immunogenic fragments of it, nucleotide sequences
CC encoding it, or inhibitor specific against it are used to manufacturing a
CC medicament for the treatment of infection due to C. pneumoniae. An
CC antibody specific against the protein can diagnose a C. pneumoniae
CC infection. C. pneumoniae infection can be prevented. Arteriosclerosis,
CC including coronary atherosclerosis, caused by C. pneumoniae can also be
CC prevented or treated. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 14 AA;
Query Match 100.0%; Score 76; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
XX |||||||
Db 1 ETPELPKPGVTPRS 14
RESULT 2
AAR94586
ID AAR94586 standard; protein; 259 AA.
XX
AC AAR94586;
XX
DT 16-OCT-2003 (revised)
DT 11-NOV-1996 (first entry)
XX
DE C. pneumoniae polypeptide antigen (polypeptide A) clone 53-3S.

XX
KM Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KM PCPN533alpha; primer; assay; detection; antibody; diagnosis; infection;
KM clone.
XX
OS Chlamydophila pneumoniae.
XX
PN WO9609320-A1.
XX
PD 28-MAR-1996.
XX
PF 20-SEP-1995; 95WO-JP001896.
XX
PR 20-SEP-1994; 94JP-00224711.
PR 28-APR-1995; 95JP-00106006.
PR 28-APR-1995; 95JP-00106008.
PR 28-APR-1995; 95JP-00106009.
PR 28-APR-1995; 95JP-00106010.
PR 28-APR-1995; 95JP-00106011.
XX
PA (HITB) HITACHI CHEM CO LTD.
XX
PI Izutsu H, Obara K, Matsumoto A;
XX
DR WPI; 1996-188399/19.
DR N-PSDB; AAT14622.
XX
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it - used for
PT detection and assay of C. pneumoniae e.g. in clinical diagnosis.
XX
PS Example 1; Page 75-77; 128pp; Japanese.
XX
CC The present sequence is the C. pneumoniae polypeptide antigen polypeptide
CC A clone, 53-3S. C. pneumoniae strain YK41 was cultured and genomic DNA
CC extracted to prep. a lambda gt11 DNA library. The library was then
CC screened with an anti-YK41 monoclonal antibody (Mab), which was prepd. by
CC fusing spleen cells from a mouse infected with YK41 with myeloma P3/NS1/1
CC -Ag4-1 to produce a Mab expressing hydridoma. The DNA obtd. was then
CC fused with the expression vector PAD431 to give PCPN533alpha. The
CC plasmid was used to transform an E. coli host, which was cultured to give
CC the antigenic polypeptide, polypeptide A. Polypeptide A and primers and
CC probes derived from its DNA can be used in assays for the detection of
CC polypeptide A antibodies and DNA, respectively, useful in the diagnosis
CC of C. pneumoniae infection. (Updated on 16-OCT-2003 to standardise OS
CC field)
XX
SQ Sequence 259 AA;
Query Match 100.0%; Score 76; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
XX |||||||
Db 164 ETPELPKPGVTPRS 177
RESULT 3
AAW01743
ID AAW01743 standard; protein; 259 AA.
XX
AC AAW01743;
XX
DT 17-OCT-2003 (revised)
DT 22-APR-1997 (first entry)
XX
DE C. pneumoniae 53 kDa antigen.
XX
KM antigen; antibody; detection; determination; epitope.
XX
OS Chlamydophila pneumoniae.
XX
PN JP08304404-A.


```
XX PD 22-NOV-1996.
XX PF 28-APR-1995; 95JP-00106014.
XX PR 28-APR-1995; 95JP-00106014.
XX PA (HITB ) HITACHI CHEM CO LTD.
XX DR WPI; 1997-056178/06.
XX DR N-PSDB; AAT59311.
XX PT Detection and determination of anti-Chlamydia pneumoniae antibody - using
XX PT the polypeptide C as the antigen.
XX PS Example 3; Page 17-18; 18pp; Japanese.
XX CC The protein is that of the Chlamydia pneumoniae 53 kDa antigen. A method
XX CC for the detection and determination of anti-C. pneumoniae antibodies in a
XX CC sample comprises using at least 5 consecutive amino acids of the
XX CC polypeptide C 73 kDa antigen (AAW01742). (Updated on 17-OCT-2003 to
XX CC standardise OS field)
SQ Sequence 259 AA;

Query Match 100.0%; Score 76; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETEPLPKPGVTPRS 14
   |||||
Db 164 ETEPLPKPGVTPRS 177

RESULT 4
AAR94580
ID AAR94580 standard; protein; 271 AA.
XX AC AAR94580;
XX DT 16-OCT-2003 (revised)
XX DT 07-NOV-1996 (first entry)
XX DE C. pneumoniae polypeptide antigen (polypeptide A) variant.
XX KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
XX KW PCPN533alpha; primer; assay; detection; antibody; diagnosis; infection;
XX KW variant.
XX OS Chlamydophila pneumoniae.
XX PN WO9609320-A1.
XX PD 28-MAR-1996.
XX PF 20-SEP-1995; 95WO-JP001896.
XX PR 20-SEP-1994; 94JP-00224711.
XX PR 28-APR-1995; 95JP-00106006.
XX PR 28-APR-1995; 95JP-00106008.
XX PR 28-APR-1995; 95JP-00106009.
XX PR 28-APR-1995; 95JP-00106010.
XX PR 28-APR-1995; 95JP-00106011.
XX PA (HITB ) HITACHI CHEM CO LTD.
XX PI Izutou H, Obara K, Matsumoto A;
XX DR WPI; 1996-188399/19.
XX DR N-PSDB; AAT14613.
XX PT Recombinant Chlamydia pneumoniae antigen and antibodies to it - used for
XX PT detection and assay of C. pneumoniae e.g. in clinical diagnosis.
```

```
XX PS Claim 2; Page 64-66; 128pp; Japanese.
XX CC The present sequence is a variant of the C. pneumoniae polypeptide
XX CC antigen, polypeptide A. C. pneumoniae strain YK41 was cultured and
XX CC genomic DNA extracted to prep. a lambda gt11 DNA library. The library was
XX CC then screened with an anti-YK41 monoclonal antibody (Mab), which was
XX CC prepd. by fusing spleen cells from a mouse infected with YK41 with
XX CC myeloma P3/NS1/1-Ag4-1 to produce a Mab expressing hybridoma. The DNA
XX CC obtd. was then fused with the expression vector pADA431 to give
XX CC PCPN533alpha. The plasmid was used to transform an E. coli host, which
XX CC was cultured to give the antigenic polypeptide, polypeptide A.
XX CC Polypeptide A and primers and probes derived from its DNA can be used in
XX CC assays for the detection of polypeptide A antibodies and DNA,
XX CC respectively, useful in the diagnosis of C. pneumoniae infection.
XX CC (Updated on 16-OCT-2003 to standardise OS field)
SQ Sequence 271 AA;

Query Match 100.0%; Score 76; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.028;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ETEPLPKPGVTPRS 14
   |||||
Db 164 ETEPLPKPGVTPRS 177

RESULT 5
AAB08323
ID AAB08323 standard; protein; 302 AA.
XX AC AAB08323;
XX DT 04-DEC-2000 (first entry)
XX DE Amino acid sequence of a Chlamydia pneumoniae His-tagged protein.
XX KW Chlamydia pneumoniae protein; infection; arteriosclerosis;
XX KW coronary atherosclerosis.
XX OS Synthetic.
XX OS Chlamydophila pneumoniae.
XX FH Key Location/Qualifiers
XX FT Cleavage-site 1..30
XX FT /note= "S-tag and thrombin cleavage site"
XX FT Misc-difference 292..302
XX FT /note= "Histidine tag"
XX PN WO200046359-A2.
XX PD 10-AUG-2000.
XX PF 28-JAN-2000; 2000WO-GB000237.
XX PR 05-FEB-1999; 99GB-00002555.
XX PA (NEUT-) NEUTEC PHARMA PLC.
XX PI Burnie JP, Matthews RC;
XX DR WPI; 2000-543485/49.
XX DR New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX DR preventing and treating C. pneumoniae infection and arteriosclerosis,
XX DR including coronary atherosclerosis.
XX PS Example; Page 31-32; 35pp; English.
XX CC The present sequence represents a His-tagged Chlamydia pneumoniae
XX CC protein. The protein, immunogenic fragments of it, nucleotide sequences
XX CC encoding it, or inhibitor specific against it are used to manufacturing a
```

CC medicament for the treatment of infection due to C. pneumoniae. An
CC antibody specific against the protein can diagnose a C. pneumoniae
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,
CC including coronary atherosclerosis, caused by C. pneumoniae can also be
CC prevented or treated

XX Sequence 302 AA;

Query Match 100.0%; Score 76; DB 3; Length 302;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPRS 14

Db 199 ETPELPKPGVTPRS 212

RESULT 6

AAR94585 AAR94585 standard; protein; 432 AA.

XX AAR94585;

XX 11-NOV-1996 (first entry)

DE DHFR/C. pneumoniae antigen variant fusion protein.

XX Polypeptide antigen; strain YK41; plasmid; probe; PCPN533T; primer;
KW assay; detection; antibody; diagnosis; infection; fusion protein;
KM dihydrofolate reductase; DHFR; variant; Chlamydia pneumoniae.

OS Synthetic.

FH Key Location/Qualifiers

FT Protein 1..160 /label= dihydrofolate reductase

FT Protein 162..432 /label= C. pneumoniae antigen

PN WO9609320-A1.

PD 28-MAR-1996.

PF 20-SEP-1995; 95WO-JP001896.

XX 20-SEP-1994; 94JP-00224711.

PR 28-APR-1995; 95JP-00106006.

PR 28-APR-1995; 95JP-00106008.

PR 28-APR-1995; 95JP-00106009.

PR 28-APR-1995; 95JP-00106010.

PR 28-APR-1995; 95JP-00106011.

XX (HITB) HITACHI CHEM CO LTD.

PI Izutsu H, Obara K, Matsumoto A;

DR WPI; 1996-188399/19.

XX N-PSDB; AAT14619.

PT Recombinant Chlamydia pneumoniae antigen and antibodies to it - used for

XX detection and assay of C. pneumoniae e.g. in clinical diagnosis.

PS Claim 23; Page 103-107; 128pp; Japanese.

CC The present sequence is the dihydrofolate reductase (DHFR) / C. pneumoniae
CC polypeptide antigen fusion protein. C. pneumoniae strain YK41 was
CC cultured and genomic DNA extracted to prep. a lambda gt11 DNA library.
CC The library was then screened with an anti-YK41 monoclonal antibody
CC (Mab), which was prepd. by fusing spleen cells from a mouse infected with
CC YK41 with myeloma P3/NS1/1-Ag4-1 to produce a Mab expressing hybridoma.
CC The DNA obtd. was then fused with DHFR DNA and the expression vector
CC PAD431 to give PCPN533T. The plasmid was used to transform an E. coli
CC host, which was cultured to give an antigenic polypeptide fusion protein.

CC The fusion protein and primers and probes derived from its DNA can be
CC used in assays for the detection of the antigenic polypeptide antibodies
CC and DNA, respectively, useful in the diagnosis of C. pneumoniae infection

XX Sequence 432 AA;

Query Match 100.0%; Score 76; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPRS 14

Db 325 ETPELPKPGVTPRS 338

RESULT 7

AAR94579 AAR94579 standard; protein; 488 AA.

XX AAR94579;

XX 16-OCT-2003 (revised)

DT 07-NOV-1996 (first entry)

DE Chlamydia pneumoniae polypeptide antigen (polypeptide A).

XX Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW PCPN533alpha; primer; assay; detection; antibody; diagnosis; infection.
OS Chlamydia pneumoniae.

FH Key Location/Qualifiers

FT Peptide 1..259 /note= "claimed N-terminal fragment"

PN WO9609320-A1.

PD 28-MAR-1996.

PF 20-SEP-1995; 95WO-JP001896.

XX 20-SEP-1994; 94JP-00224711.

PR 28-APR-1995; 95JP-00106006.

PR 28-APR-1995; 95JP-00106008.

PR 28-APR-1995; 95JP-00106009.

PR 28-APR-1995; 95JP-00106010.

PR 28-APR-1995; 95JP-00106011.

XX (HITB) HITACHI CHEM CO LTD.

PI Izutsu H, Obara K, Matsumoto A;

DR WPI; 1996-188399/19.

XX N-PSDB; AAT14612.

PT Recombinant Chlamydia pneumoniae antigen and antibodies to it - used for

XX detection and assay of C. pneumoniae e.g. in clinical diagnosis.

PS Claim 1; Page 60-64; 128pp; Japanese.

CC The present sequence is the C. pneumoniae polypeptide antigen,
CC polypeptide A. C. pneumoniae strain YK41 was cultured and genomic DNA
CC extracted to prep. a lambda gt11 DNA library. The library was then
CC screened with an anti-YK41 monoclonal antibody (Mab), which was prepd. by
CC fusing spleen cells from a mouse infected with YK41 with myeloma P3/NS1/1
CC -Ag4-1 to produce a Mab expressing hybridoma. The DNA obtd. was then
CC fused with the expression vector PAD431 to give PCPN533alpha. The
CC plasmid was used to transform an E. coli host, which was cultured to give
CC the antigenic polypeptide, polypeptide A. Polypeptide A and primers and
CC probes derived from its DNA can be used in assays for the detection of
CC polypeptide A antibodies and DNA, respectively, useful in the diagnosis
CC of C. pneumoniae infection. (Updated on 16-OCT-2003 to standardise OS
CC field)

```
XX Sequence 488 AA;
SQ
Query Match 100.0%; Score 76; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
Db 164 ETPELPKPGVTPRS 177
RESULT 8
ABO23550
ID ABO23550 standard; protein; 493 AA.
XX
AC ABO23550;
XX
DT 04-SEP-2003 (first entry)
XX
DE Chlamydophila pneumoniae CWL029 outlier protein #4.
XX
KW Candidate protein identification; pathogen; anti-infective;
KW outlier protein; virulence protein; antigen; drug target protein;
KW pathogenic organism; antimicrobial.
XX
OS Chlamydophila pneumoniae CWL029.
XX
PN US2003039963-A1.
XX
PD 27-FEB-2003.
XX
PF 30-MAR-2001; 2001US-00820843.
XX
PR 30-MAR-2001; 2001US-00820843.
XX
PA (BRAH/) BRAHMACHARI S K.
PA (RAMA/) RAMACHANDRAN S.
PA (NAND/) NANDI T.
PA (BHIM/) BHIMARAO C.
XX
PI Brahmachari SK, Ramachandran S, Nandi T, Bhimarao C;
XX
DR WPI; 2003-492159/46.
XX
PT Identifying candidate proteins useful as anti-infectives involves
PT matching outlier protein sequences with protein sequences in databases.
XX
PS Example 7; Page 54-55; 117pp; English.
XX
CC The present invention relates to a method for identifying candidate
CC proteins in pathogens useful as anti-infectives. The invention discloses
CC a computational method which involves the calculation of several sequence
CC attributes and their subsequence analysis results in the identification
CC of outlier proteins in different pathogens. The method is useful for the
CC identification of outlier proteins (e.g. virulence proteins, antigens or
CC proteins used as drug targets) in pathogenic organisms. The method of the
CC invention provides reproducible results as it does not depend on the
CC variable biochemical characterisation of proteins. ABO23500-ABO23617
CC represent outlier proteins identified from different pathogenic organisms
XX
SQ Sequence 493 AA;
```

RESULT 9

```
AAV35449
ID AAV35449 standard; protein; 496 AA.
XX
AC AAV35449;
XX
DT 17-OCT-2003 (revised)
DT 13-SEP-1999 (first entry)
XX
DE Amino acid sequence of a Chlamydia pneumoniae protein.
XX
KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine;
KW neutralising epitope.
XX
OS Chlamydophila pneumoniae.
XX
PN WO927105-A2.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-IB001890.
XX
PR 21-NOV-1997; 97FR-00014673.
PR 04-NOV-1998; 98US-0107078P.
XX
PA (GEST ) GENSET.
XX
PI Griffiths R;
XX
DR WPI; 1999-357842/30.
XX
PT Genome sequence of Chlamydia pneumoniae.
XX
PS Page 1223-1224; Disclosure; 1912pp; English.
XX
CC AAY34584-Y35879 represent the proteins encoded by all the open reading
CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C.
CC pneumoniae causes respiratory disease such as pneumonia and bronchitis
CC and is thought to be a contributing factor in heart disease, sarcoidosis,
CC sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The
CC polypeptides encoded by the open reading frames of the C. pneumoniae
CC genome (see AAY34584-Y35879) can be used in immunogenic compositions as
CC vaccines. Vectors containing C. pneumoniae nucleotides sequences can also
CC be used as immunogenic compositions, especially where the vector directs
CC the expression of a neutralising epitope of C. pneumoniae. (Updated on 17
CC -OCT-2003 to standardise OS field)
XX
SQ Sequence 496 AA;
```

```
Query Match 100.0%; Score 76; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ETPELPKPGVTPRS 14
Db 167 ETPELPKPGVTPRS 180
RESULT 10
AAB08322
ID AAB08322 standard; protein; 496 AA.
XX
AC AAB08322;
XX
DT 12-SEP-2003 (revised)
DT 04-DEC-2000 (first entry)
XX
DE Amino acid sequence of a Chlamydia pneumoniae protein.
XX
KW Chlamydia pneumoniae protein; infection; arteriosclerosis;
KW coronary atherosclerosis.
XX
OS Chlamydocphila pneumoniae.
```

XX WO200046359-A2.
PN 10-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-GB000237.
PF 05-FEB-1999; 99GB-00002555.
PR
XX (NEUT-) NEUTEC PHARMA PLC.
PA
XX Burnie JP, Matthews RC;
PI
XX WPI; 2000-543485/49.
DR N-PSDB; AAA63621.
DR
XX
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
PT preventing and treating C. pneumoniae infection and atherosclerosis,
PT including coronary atherosclerosis.
XX
XX Claim 1; Page 30-31; 35pp; English.
XX
CC The present sequence represents a Chlamydia pneumoniae protein. The
CC protein, immunogenic fragments of it, nucleotide sequences encoding it,
CC or inhibitor specific against it are used to manufacturing a medicament
CC for the treatment of infection due to C. pneumoniae. An antibody specific
CC against the protein can diagnose a C. pneumoniae infection. C. pneumoniae
CC infection can be prevented. Atherosclerosis, including coronary
CC atherosclerosis, caused by C. pneumoniae can also be prevented or
CC treated. (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 496 AA;

Query Match 100.0%; Score 76; DB 3; Length 496;
Best Local Similarity 100.0%; Pred. No. 0.048;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
|||
Db 167 ETPELPKPGVTPRS 180

RESULT 11
AAR94584
ID AAR94584 standard; protein; 649 AA.
XX
AC AAR94584;
XX
DT 11-NOV-1996 (first entry)
XX
DE DHFR/C. pneumoniae antigen fusion protein (polypeptide B).
XX
KW Polypeptide antigen; polypeptide B; strain YK41; plasmid; probe;
KW PCPN533T; primer; assay; detection; antibody; diagnosis; infection;
KW fusion protein; dihydrofolate reductase; DHFR; Chlamydia pneumoniae.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Protein 1..160
FT /label= dihydrofolate reductase
FT Protein 162..649
FT /label= C. pneumoniae antigen
XX
PN WO9609320-A1.
XX
PD 28-MAR-1996.
XX
PF 20-SEP-1995; 95WO-JP001896.
XX
PR 20-SEP-1994; 94JP-00224711.
PR 28-APR-1995; 95JP-00106006.
PR 28-APR-1995; 95JP-00106008.

PR 28-APR-1995; 95JP-00106009.
PR 28-APR-1995; 95JP-00106010.
PR 28-APR-1995; 95JP-00106011.
XX
XX (HITB) HITACHI CHEM CO LTD.
PA
XX Izutsu H, Obara K, Matsumoto A;
PI
XX WPI; 1996-188399/19.
DR N-PSDB; AAT14618.
DR
XX
XX Recombinant Chlamydia pneumoniae antigen and antibodies to it - used for
PT detection and assay of C. pneumoniae e.g. in clinical diagnosis.
PT
XX
PS Claim 22; Page 97-103; 128pp; Japanese.
XX
XX The present sequence is the dihydrofolate reductase (DHFR)/ C. pneumoniae
CC polypeptide antigen fusion protein, polypeptide B. C. pneumoniae strain
CC YK41 was cultured and genomic DNA extracted to prep. a lambda gt11 DNA
CC library. The library was then screened with an anti-YK41 monoclonal
CC antibody (MAb), which was prepd. by fusing spleen cells from a mouse
CC infected with YK41 with myeloma P3/NS1/1-Ag4-1 to produce a MAb
CC expressing hybridoma. The DNA obtd. was then fused with DHFR DNA and the
CC expression vector PAD431 to give PCPN533T. The plasmid was used to
CC transform an E. coli host, which was cultured to give the antigenic
CC polypeptide fusion protein, polypeptide B. Polypeptide B and primers and
CC probes derived from its DNA can be used in assays for the detection of
CC antigenic polypeptide antibodies and DNA, respectively, useful in the
CC diagnosis of C. pneumoniae infection
XX
SQ Sequence 649 AA;

Query Match 100.0%; Score 76; DB 2; Length 649;
Best Local Similarity 100.0%; Pred. No. 0.061;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTPRS 14
|||
Db 325 ETPELPKPGVTPRS 338

RESULT 12
AAV37632
ID AAV37632 standard; protein; 481 AA.
XX
AC AAV37632;
XX
DT 07-OCT-1999 (first entry)
XX
DE Protein which is specific to Chlamydia trachomatis.
XX
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.
XX
OS Chlamydia trachomatis.
XX
PN WO9928475-A2.
XX
PD 10-JUN-1999.
XX
PF 27-NOV-1998; 98WO-IB001939.
XX
PR 28-NOV-1997; 97FR-00015041.
PR 17-DEC-1997; 97FR-00016034.
PR 04-NOV-1998; 98US-0107077P.
XX
PA (GEST) GENSET.
XX
PI Griffiths R;
XX
DR WPI; 1999-371125/31.

XX Genome sequence of Chlamydia trachomatis.
PT
XX
PS Disclosure; Page 1271-1272; 1755bp; English.
XX
CC AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
CC of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
CC can also be used to control growth of the microorganism. Chlamydia
CC trachomatis is responsible for a large number of diseases, e.g. eye
CC diseases such as conventional trachoma, nonendemic trachoma,
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC nongonococcal urethritis, epidymitis, cervicitis, salpingitis,
CC perihepatitis, Bartholin's; pneumopathy in breast feeding infants; and
CC venereal lymphogranulomatosis. The polypeptides of the invention may be
CC of use in treating these diseases
XX
SQ Sequence 481 AA;

Query Match 72.4%; Score 55; DB 2; Length 481;
Best Local Similarity 76.9%; Pred. No. 29;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPR 13
||| ||| |||
Db 164 ETPGLPKPSTTPR 176

RESULT 13
ABU27094
ID ABU27094 standard; protein; 487 AA.
XX
AC ABU27094;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #12621.
XX
KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Chlamydia trachomatis.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI: 2003-029926/02.
DR N-PSDB; ACA30964.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 55018; 1766bp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 487 AA;

Query Match 72.4%; Score 55; DB 6; Length 487;
Best Local Similarity 76.9%; Pred. No. 29;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTPR 13
||| ||| |||
Db 161 ETPGLPKPSTTPR 173

RESULT 14
ABG03392
ID ABG03392 standard; protein; 535 AA.
XX
AC ABG03392;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3383.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR N-PSDB; AAS67579.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 33751; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 535 AA;

Query Match 65.8%; Score 50; DB 4; Length 535;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTPRS 14
:|||:|||||
Db 178 QTPSVPKPGLEPTS 191

RESULT 15
ABG15076
ID ABG15076 standard; protein; 228 AA.
XX AC ABG15076;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #15067.
XX KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-6393362/73.
XX DR N-PSDB; AAS79263.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
PS Claim 20; SEQ ID NO 45435; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 228 AA;

Query Match 61.8%; Score 47; DB 4; Length 228;
Best Local Similarity 66.7%; Pred. No. 1.7e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
:|||:|||||
Db 28 QTPELQRPGVTP 39

RESULT 16
AAV17406
ID AAV17406 standard; protein; 1012 AA.
XX AC AAV17406;
XX DT 22-JUL-1999 (first entry)
XX DE Human atrophin-1 related protein.
XX KM Human; atrophin-1 related protein; Charcot-Marie-Tooth disease;
XX KW Schwartz-Jampel syndrome; crystalline corneal dystrophy;
XX KW dentatorubral pallidolysian atrophy; ataxia.
XX OS Homo sapiens.
XX PN WO9921983-A1.
XX PD 06-MAY-1999.
XX PF 27-OCT-1997; 97WO-CN000108.
XX PR 27-OCT-1997; 97WO-CN000108.
XX PA (UYHU-) UNIV HUNAN MEDICAL.
XX PI Xia J, Liu C, Ruan Q, Wang D, Deng H;
XX DR WPI; 1999-326701/27.
XX DR N-PSDB; AAX56436.
XX PT Human atrophin-1 related gene and polypeptide useful for treating Charcot
XX PT -Marie-Tooth disease and Schwartz-Jampel syndrome.
XX PS Claim 11; Page 25-27; 34pp; English.
XX CC The present sequence represents human atrophin-1 related protein. The
XX CC atrophin-1 related gene (I), its expressed polypeptides (II) and agonists
XX CC of (II) are used to treat or prevent conditions requiring an increase in

CC activity and/or expression of (II), while antagonists of (II), inhibitory
CC nucleic acid and/or competitive polypeptides are used to treat conditions
CC requiring reduced activity or expression of (II). Typical of these
CC conditions are Charcot-Marie-Tooth disease; Schwartz-Jampel syndrome;
CC crystalline corneal dystrophy, dentatorubral pallidolysian atrophy and
CC ataxia. These diseases, or susceptibility to them, can be diagnosed by
CC detecting mutations in (II)-encoding genes and/or by measuring levels of
CC (II). Cells that express (II), or their membranes, can be used to screen
CC for specific (ant)agonists, potential therapeutic agents. (I), or its
CC fragments, are used as probes to isolate related sequences; as reagents
CC for research, diagnosis and drug screening; for chromosome identification
CC ; and for expressing recombinant (II). (II) and their fragments, are used
CC to raise specific antibodies (useful for isolation and identification of
CC (II)-expressing clones, for purification of (II), as therapeutic
CC antagonists, as assay reagents and in drug screening); in vaccines; to
CC screen for (ant)agonists and to identify specific receptors
XX
SQ Sequence 1012 AA;

Query Match 61.8%; Score 47; DB 2; Length 1012;
Best Local Similarity 72.7%; Pred. No. 6.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TPBLPKPGVTP 12
Db 204 TPQLPTPGPTP 214

RESULT 17
ABR59716
ID ABR59716 standard; protein; 1566 AA.
XX ABR59716;
AC ABR59716;
XX 25-JUL-2003 (first entry)
DT 25-JUL-2003 (first entry)
XX Human RERE.
DE Human RERE.
XX
XX Human; T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;
KW immunosuppressive; antiaesthetic; antiallergic; antiinflammatory;
KW lymphocyte activation; lymphocyte migration; cytokine production;
KW cell surface marker expression; antibody production; apoptosis; allergy;
KW antibody proliferation; antibody differentiation; hypersensitivity;
KW graft versus host disease; inflammation; RERE.
XX
XX Homo sapiens.
OS
XX
XX WO2003029277-A2.
PN
XX 10-APR-2003.
PD
XX 02-OCT-2002; 2002WO-US031618.
PF
XX 03-OCT-2001; 2001US-0327212P.
PR
XX (RIGE-) RIGEL PHARM INC.
PA
XX Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;
PI WPI; 2003-363276/34.
XX N-PSDB; ACC81113.
DR
XX Identifying a compound that modulates T lymphocyte activation, useful for
PT monitoring changes in cell surface marker expression, comprises
PT contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with
PT a compound.
XX
XX
PS Disclosure; Page 93; 126pp; English.
XX
XX The invention relates to a novel method for identifying a compound that
CC modulates T lymphocyte activation. The method comprises contacting a T
CC cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,
CC where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic

CC acid that hybridises to a nucleic acid encoding a polypeptide having a
CC sequence selected from two 606-amino acid sequence and a 415-amino acid
CC sequence given in the specification. The method of the invention has
CC immunosuppressive, antiaesthetic, antiallergic, and antiinflammatory
CC activity. The method is useful for identifying compounds that modulate
CC lymphocyte activation and migration, and for monitoring changes in cell
CC surface marker expression, cytokine production, antibody production,
CC proliferation and differentiation, and apoptosis, using either cell lines
CC or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as
CC drug targets for compounds that suppress or activate lymphocyte
CC activation and migration, e.g. for the treatment of diseases in which
CC modulation of the immune response is desired such as delayed type
CC hypersensitivity reactions, asthma, allergies, graft versus host disease,
CC and acute and chronic inflammation. Modulators of lymphocyte activation
CC are useful for treating disorders related T and B cell activation and
CC migration. The present sequence is used in the exemplification of the
CC invention
XX
SQ Sequence 1566 AA;

Query Match 61.8%; Score 47; DB 6; Length 1566;
Best Local Similarity 72.7%; Pred. No. 9.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 TPBLPKPGVTP 12
Db 758 TPQLPTPGPTP 768

RESULT 18
ABG20603
ID ABG20603 standard; protein; 133 AA.
XX ABG20603;
AC ABG20603;
XX 13-FEB-2002 (first entry)
DT 13-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #20594.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS84790.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 50962; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 60.5%; Score 46; DB 4; Length 133;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
:||||:||||
Db 52 QTPELQPGVPP 63

RESULT 19
ABG23946
ID ABG23946 standard; protein; 133 AA.
XX
AC ABG23946;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #23937.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS88133.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 54305; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 133 AA;

Query Match 60.5%; Score 46; DB 4; Length 133;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
:||||:||||
Db 52 QTPELQPGVPP 63

RESULT 20
ABG09708
ID ABG09708 standard; protein; 181 AA.
XX
AC ABG09708;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9699.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS73895.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 40067; 103bp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging

CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 181 AA;

Query Match 60.5%; Score 46; DB 4; Length 181;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKRGVTP 12
 :|||||:|||||
Db 100 QTPELQQPGVPP 111

RESULT 21
ABG06822
ID ABG06822 standard; protein; 292 AA.
XX
AC ABG06822;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6813.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS71009.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 37181; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The

CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 292 AA;

Query Match 60.5%; Score 46; DB 4; Length 292;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKRGVTP 12
 :|||||:|||||
Db 211 QTPELQQPGVPP 222

RESULT 22
ABG20398
ID ABG20398 standard; protein; 294 AA.
XX
AC ABG20398;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20389.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS84585.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 50757; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 294 AA;

Query Match 60.5%; Score 46; DB 4; Length 294;
Best Local Similarity 66.7%; Pred. No. 2.9e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
:||||:||||
Db 213 QTPELQPGVPP 224

RESULT 23

ABG20590
ID ABG20590 standard; protein; 309 AA.

XX AC ABG20590;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #20581.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KM food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS84777.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PS Claim 20; SEQ ID NO 50949; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 309 AA;

Query Match 60.5%; Score 46; DB 4; Length 309;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETPELPKPGVTP 12
:||||:||||
Db 228 QTPELQPGVPP 239

RESULT 24

ABG09703
ID ABG09703 standard; protein; 309 AA.

XX AC ABG09703;

XX DT 13-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #9694.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KM food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS73890.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.

XX PS Claim 20; SEQ ID NO 40062; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SO Sequence 309 AA;
Query Match 60.5%; Score 46; DB 4; Length 309;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETEPLPKPGVTP 12
:||||:||||
Db 228 QTEPLQQPGVVP 239

RESULT 25
ABG26564
ID ABG26564 standard; protein, 310 AA.
XX
AC ABG26564;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #26555.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS90751.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 56923; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
XX
SO Sequence 310 AA;

Query Match 60.5%; Score 46; DB 4; Length 310;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ETEPLPKPGVTP 12
:||||:||||
Db 229 QTEPLQQPGVVP 240

RESULT 26
ABG09701
ID ABG09701 standard; protein, 310 AA.
XX
AC ABG09701;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #9692.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS73888.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 40060; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 310 AA;
Query Match 60.5%; Score 46; DB 4; Length 310;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTP 12
:|||||:|||||
Db 229 QTPELQQPGVPP 240

RESULT 27
ABG00825
ID ABG00825 standard; protein; 310 AA.
XX
AC ABG00825;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #816.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS65012.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 31184; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 310 AA;

Query Match 60.5%; Score 46; DB 4; Length 310;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches .8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTP 12
:|||||:|||||
Db 229 QTPELQQPGVPP 240

RESULT 28
ABG20589
ID ABG20589 standard; protein; 310 AA.
XX
AC ABG20589;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20580.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS84776.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 50948; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 310 AA;
Query Match 60.5%; Score 46; DB 4; Length 310;
Best Local Similarity 66.7%; Pred. No. 3e+02;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTP 12
:||||:|||||
Db 229 QTEPQQPGVPP 240

RESULT 29
ABG18606
ID ABG18606 standard; protein; 310 AA.
XX
AC ABG18606;
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #18597.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS82793.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 48965; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 310 AA;

Query Match 60.5%; Score 46; DB 4; Length 310;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
:||||:|||||
Db 229 QTEPQQPGVPP 240

RESULT 30
ABG03062
ID ABG03062 standard; protein; 310 AA.
XX
AC ABG03062;
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #3053.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS67249.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 33421; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 310 AA;

Query Match 60.5%; Score 46; DB 4; Length 310;
Best Local Similarity 66.7%; Pred. No. 3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ETPELPKPGVTP 12
:||||:|||||

Db 229 QTPELQPGVPP 240

RESULT 31
ABB63260
ID ABB63260 standard; protein; 485 AA.
XX
AC ABB63260;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 16572.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li FWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL07363.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 16572; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 485 AA;

Query Match 60.5%; Score 46; DB 4; Length 485;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
Db 438 PELPKPKLTP 447

RESULT 32
AAG80006
ID AAG80006 standard; protein; 485 AA.
XX
AC AAG80006;
XX
DT 14-JAN-2002 (first entry)
XX
DE D. melanogaster ligand-controlled anion channel protein SEQ ID 12.
XX
KW ligand-controlled anion channel; fruitfly; invertebrate; screening;

KW plant-protection agent; insecticide; transgenic.
XX
OS Drosophila melanogaster.
XX
PN EP1136502-A1.
XX
PD 26-SEP-2001.
XX
PF 06-MAR-2001; 2001EP-00104577.
XX
PR 18-MAR-2000; 2000DE-01013619.
XX
PA (FARB) BAYER AG.
XX
PI Franken E, Friedrich G, Raming K;
XX
DR WPI; 2001-598685/68.
DR N-PSDB; AAI68567.
XX
PT New polypeptide subunit of invertebrate ligand-controlled anion channel,
PT useful in screening for potential plant-protection agents, especially
PT insecticides.
XX
PS Claim 2; Page 36-38; 60pp; German.
XX
CC This invention describes novel polypeptide subunit(s) of invertebrate
CC ligand-controlled anion channels which are useful in screening for
CC potential plant-protection agents and have insecticidal activity. The
CC invention also describes the production of DNA constructs, vectors, host
CC cells, antibodies and transgenic invertebrates which are also used in the
CC identification of plant protection agents which alter the expression of
CC the ligand controlled anion channel protein. The products of the
CC invention are also used to identify genes that encode polypeptides
CC involved in assembly of similar ligand-controlled channels in insects.
CC The novel polypeptides described in the specification are suitable for
CC use in high throughput assays. This sequence represents a ligand-
CC controlled anion channel from the fruitfly Drosophila melanogaster
XX
SQ Sequence 485 AA;

Query Match 60.5%; Score 46; DB 4; Length 485;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
Db 438 PELPKPKLTP 447

RESULT 33
AAG67566
ID AAG67566 standard; protein; 485 AA.
XX
AC AAG67566;
XX
DT 26-NOV-2001 (first entry)
XX
DE Amino acid sequence of Drosophila LGIC clone AC15-4/AC15-4.
XX
KW Ligand-gated ion channel; LGIC; AC05-10; AC05-11; AC15-4; AC15-25;
KW glutamate-related channel; insecticide; miticide; nematocide;
KW crop protection.
XX
OS Drosophila melanogaster.
XX
PN WO200164705-A1.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US006096.
XX
PR 02-MAR-2000; 2000US-0186645P.
XX

PA (MERI) MERCK & CO INC.
XX
PI Cully DF, Hirschberg B, Yuan J, Zheng Y;
XX
DR WPI; 2001-570677/64.
DR N-PSDB; AAH78309, AAH78310.
XX
PT New Drosophila ligand-gated ion channel protein, for identifying ion
PT channel protein modulators which act as insecticidal, mitacidal and/or
PT nematocidal agents against worms, fleas, ticks and lice.
XX
PS Claim 26; Fig 7; 85pp; English.
XX
CC The present sequence represents a Drosophila melanogaster ligand-gated
CC ion channel (LGIC) protein. The specification describes LGIC proteins
CC designated LGIC AC05-10, AC05-11 and AC15-4/AC15-25. LGIC polynucleotides
CC are useful for identifying a compound that modulates glutamate-related
CC channel protein activity. The LGIC protein forms a forms homomultimer or
CC heteromultimer channel receptors which provide additional screening
CC targets to identify modulators, which may act as effective insecticidal,
CC mitacidal and/or nematocidal treatment (i.e., active against parasitic,
CC invertebrate species such as worms, fleas, tick and lice) for use in
CC human health and/or crop protection. LGIC is useful as a novel
CC insecticide target
XX
SQ Sequence 485 AA;

Query Match 60.5%; Score 46; DB 4; Length 485;
Best Local Similarity 80.0%; Pred. No. 4.5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
||| ||| : |||
Db 438 PELPKPKLTP 447

RESULT 34
AAG80005
ID AAG80005 standard; protein; 535 AA.
XX
AC AAG80005;
XX
DT 14-JAN-2002 (first entry)
XX
DE D. melanogaster ligand-controlled anion channel protein SEQ ID 10.
XX
KW ligand-controlled anion channel; fruitfly; invertebrate; screening;
KW plant-protection agent; insecticide; transgenic.
XX
OS Drosophila melanogaster.
XX
PN EP1136502-A1.
XX
PD 26-SEP-2001.
XX
PF 06-MAR-2001; 2001EP-00104577.
XX
PR 18-MAR-2000; 2000DE-01013619.
XX
PA (FARB) BAYER AG.
XX
PI Franken E, Friedrich G, Raming K;
XX
DR WPI; 2001-598685/68.
DR N-PSDB; AAI68566.
XX
PT New polypeptide subunit of invertebrate ligand-controlled anion channel,
PT useful in screening for potential plant-protection agents, especially
PT insecticides.
XX
PS Claim 2; Page 31-33; 60pp; German.
XX
CC This invention describes novel polypeptide subunit(s) of invertebrate

CC ligand-controlled anion channels which are useful in screening for
CC potential plant-protection agents and have insecticidal activity. The
CC invention also describes the production of DNA constructs, vectors, host
CC cells, antibodies and transgenic invertebrates which are also used in the
CC identification of plant protection agents which alter the expression of
CC the ligand controlled anion channel protein. The products of the
CC invention are also used to identify genes that encode polypeptides
CC involved in assembly of similar ligand-controlled channels in insects.
CC The novel polypeptides described in the specification are suitable for
CC use in high throughput assays. This sequence represents a ligand-
CC controlled anion channel from the fruitfly Drosophila melanogaster
XX
SQ Sequence 535 AA;

Query Match 60.5%; Score 46; DB 4; Length 535;
Best Local Similarity 80.0%; Pred. No. 5e+02;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PELPKPGVTP 12
||| ||| : |||
Db 438 PELPKPKLTP 447

RESULT 35
AAW56444
ID AAW56444 standard; protein; 2963 AA.
XX
AC AAW56444;
XX

DT 27-AUG-2003 (revised)
DT 31-JUL-1998 (first entry)
XX
DE Fragment HGJ1789 of a new Hepatitis virus g protein.
XX
KW Hepatitis g virus gene; diagnosis; treatment; Hepatitis g virus disease.
XX
OS Hepatitis G virus.
XX

FH Key Location/Qualifiers
FT Misc-difference 648. .667
FT /note="nucleotides encoding these amino acids not given"
XX
PN JP10108685-A.
XX
PD 28-APR-1998.
XX

PF 10-AUG-1997; 97JP-00227387.
XX
PR 10-AUG-1996; 96JP-00227639.
XX
PA (BMLB-) BML KK.
XX

DR WPI; 1998-304974/27.
DR N-PSDB; AAV23078.
XX
PT New hepatitis G virus gene - useful for diagnosing and treating diseases
PT caused by virus.
XX
PS Claim 1; Page 101-104; 128pp; Japanese.
XX

CC The present sequence represents a fragment of a new Hepatitis g virus
CC protein (see also AAW56441-50 for other fragments). RNA was synthesised
CC from the serum of nine patients judged positive for Hepatitis g virus and
CC cDNA synthesised from this RNA. The cDNA was used as a template in
CC several PCR reactions to isolate fragments of the new gene. The gene may
CC be useful for diagnosing and developing treatments for Hepatitis g virus
CC diseases. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 2963 AA;

Query Match 60.5%; Score 46; DB 2; Length 2963;
Best Local Similarity 66.7%; Pred. No. 2.3e+03;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ETPELPKPGVTP 12
| | | | |
Db 2879 EVPALPAPGVSP 2890

RESULT 36
ID ABG09669 standard; protein; 95 AA.
AC ABG09669;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #9660.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS73856.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 40028; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 95 AA;

Query Match 59.2%; Score 45; DB 4; Length 95;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ETPELPKPGVTP 12

Db 8 KTEPLOQPGVPP 19
: | | | : | | | |
| | | | |

RESULT 37
ID ABG08703 standard; protein; 158 AA.
AC ABG08703;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #8694.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS72890.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 39062; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 158 AA;

Query Match 59.2%; Score 45; DB 4; Length 158;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ETPELPKPGVTP 12

Db 71 KTEPLOQPGVPP 82

RESULT 38

ABG07541

ID ABG07541 standard; protein; 200 AA.

XX AC ABG07541;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #7532.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US008631.

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS71728.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

PS Claim 20; SEQ ID NO 37900; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 200 AA;

Query Match 57.9%; Score 44; DB 4; Length 200;

Best Local Similarity 66.7%; Pred. No. 3.8e+02;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 PELPKPGVTPRS 14

Db 172 PEKNKPGATPRT 183

RESULT 39

ADJ11540

ID ADJ11540 standard; protein; 241 AA.

XX AC ADJ11540;

DT 20-MAY-2004 (first entry)

DE Rice protein modulated by post-transcriptional gene silencing SegID 176.

KM rice; post-transcriptional gene silencing; PTGS; plant; trans-activation;
KW cereal; plant-viral interaction.

OS Oryza sp.

PN US2003135888-A1.

PD 17-JUL-2003.

PF 26-SEP-2002; 2002US-00259165.

PR 26-SEP-2001; 2001US-0325277P.

PR 27-MAR-2002; 2002US-0368327P.

PR 04-APR-2002; 2002US-0370620P.

XX (ZHUT/) ZHU T.

PA (WANG/) WANG X.

PA (CHAN/) CHANG H.

PA (BRIG/) BRIGGS S P.

PA (COOP/) COOPER B.

PA (GLAZ/) GLAZEBROOK J.

PA (GOLF/) GOLF S A.

PA (KATA/) KATAGIRI F.

PA (KREP/) KREPS J.

PA (MOUG/) MOUGHAMER T.

PA (PROV/) PROVART N.

PA (RICK/) RICHE D.

PI Zhu T, Wang X, Chang H, Briggs SP, Cooper B, Glazebrook J;

PI Goff SA, Katagiri F, Kreps J, Moughamer T, Provart N, Riche D;

DR WPI; 2003-829655/77.

DR N-PSDB; ADJ11539.

PT New polynucleotide, useful for modulating gene expression within a cell
PT by posttranscriptional gene silencing.

PS Claim 7; SEQ ID NO 176; 79pp; English.

CC This invention relates to a novel method for identifying isolated
CC polynucleotides that are modulated by post-transcriptional gene silencing
CC (PTGS). Specifically, it refers to the regulation of gene expression in
CC plants via PTGS and the trans-activation of homologous genes due to
CC increased RNA degradation. The present invention describes clusters of
CC polynucleotides from cereals, in particular rice, as well as homologues
CC and the polypeptide sequences derived thereof, where gene expression is
CC altered in response to PTGS. As such, the elucidation of gene silencing
CC mechanisms can lead to more efficiently expressed transgenes, and can
CC also improve the understanding of plant-viral interactions and targeting
CC the suppression of specific plant genes. This polypeptide sequence is a
CC rice protein sequence that is modulated by gene silencing, given in an
CC exemplification of the invention. NOTE: This sequence does not appear in
CC the printed specification but has been obtained in electronic format from
CC the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID=20030135888.

XX Sequence 241 AA;

Query Match 57.9%; Score 44; DB 7; Length 241;

Best Local Similarity 70.0%; Pred. No. 4.5e+02;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 PELPKPGVTP 12

Db 147 PELPKPTISP 156

```
RESULT 40
ADA54280
ID ADA54280 standard; protein; 299 AA.
XX
AC ADA54280;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 1848.
XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW Inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Ibogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA52641.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 1848; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 299 AA;

Query Match 57.9%; Score 44; DB 6; length 299;
Best Local Similarity 58.3%; Pred. No. 5.4e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ETPELPKPGVTP 12
Db 281 ESPEHPRPGLPP 292
```

Search completed: October 28, 2005, 22:11:55
Job time : 162 secs